

AD-A066 139

CALIFORNIA UNIV DAVIS SIGNAL AND IMAGE PROCESSING LAB F/G 9/3  
EXTRAPOLATION AND SPECTRAL ESTIMATION TECHNIQUES FOR DISCRETE T--ETC(U)  
NOV 78 A K JAIN, S RANGANATH DAAG29-78-A-0206

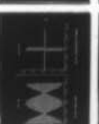
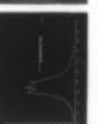
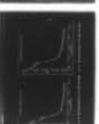
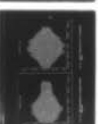
UNCLASSIFIED

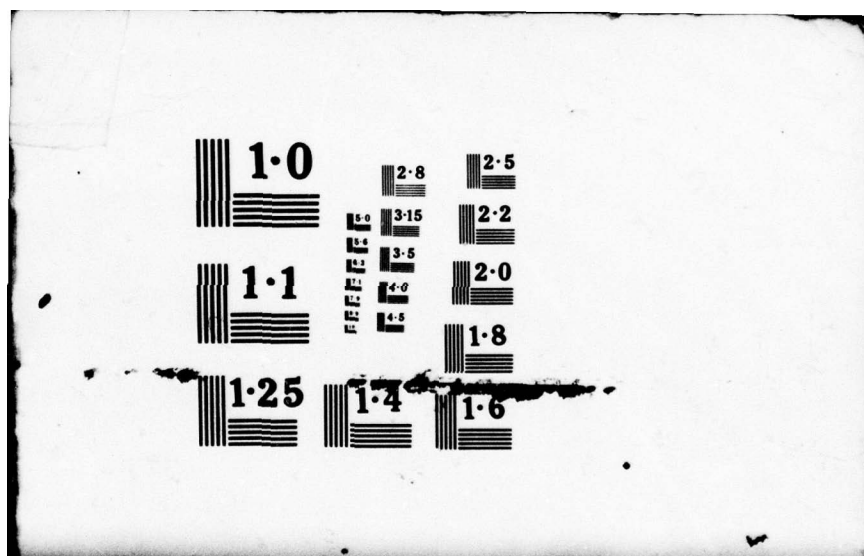
SIPL-78-1

ARO-16222.1-EL

NL

1 OF 2  
ADA  
066139





AD A0 661 39

UCD

University of California, Davis



DDC FILE COPY

SIGNAL & IMAGE PROCESSING LABORATORY

TECHNICAL REPORT SIPL-78-1

19 03 19 023

12

SIGNAL AND IMAGE PROCESSING LABORATORY

9 TECHNICAL REPORT,

14 SIPL-78-1

Research supported in part by the Army Research Office under grant No. DAAG29-78-A-0206 and in part by Rome Air Development Center under contract F30602-75-C-0122.

15 ✓ DAAG 29-78-A-4206

DDC  
RECEIVED  
MAR 21 1979  
C

18 ARO

19 16222.1-EB

6 EXTRAPOLATION AND SPECTRAL ESTIMATION TECHNIQUES  
FOR DISCRETE TIME SIGNALS.

10 Anil K. Jain  
Surendra/Ranganath

11 Nov 1978

12 145 p.

411 110

79 03 79 025

mt

**SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)**

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

# ACKNOWLEDGMENT

The work reported here was performed in part at the State University of New York at Buffalo under a contract from Rome Air Development Center, Griffiss Air Base, Rome, New York. It is a pleasure to acknowledge the support of RADC and in particular the efforts of Mr. Paul Van Etten in various phases of this project. A portion of this work was performed under the sponsorship of Army Research Office, Durham, North Carolina, under grant No. DAAG29-78-A-0206.

ACCESSION for	
NTIS	White Section <input checked="" type="checkbox"/>
DOC	B.H. Section <input type="checkbox"/>
UNCLASSIFIED	
CLASSIFICATION	
DISTRIBUTION AND AVAILABILITY	
SPECIAL	
A	

# ABSTRACT

This report considers spectral estimation and extrapolation techniques for discrete time, band limited signals, (i.e., signals whose bandwidth is less than  $\frac{1}{T}$  cycles; if  $T$  sec. is the sampling interval) which are observable only for a finite duration. The objective is to determine the spectrum (or power spectrum) of these signals. It is shown that the estimated spectrum can be improved considerably (over a periodogram or Maximum entropy spectrum) by first extrapolating the given observations beyond the observation interval. Also, we consider the problem of extrapolation of signal in the presence of noise or other interfering signals.

Several new results and algorithms are presented. First, it is shown that some of the existing extrapolation methods for continuous signals when extended to sampled data do not converge to the exact original time-unlimited signal. Rather, one only expects to get a minimum norm least squares estimate. And, we find that Papoulis' [8] iterative extrapolation algorithm is a special case of a gradient algorithm with linear convergence. It is shown that an infinite extrapolation matrix introduced in [10] does not exist and is ill-conditioned at best when approximated to a finite matrix. The new extrapolation algorithms include a discrete prolate spheroidal wave function (PSWF) expansion, a conjugate gradient iterative algorithm, a mean square extrapolation filter and a recursive Kalman filter type extrapolator. The latter two algorithms also consider the noise statistics in extrapolation of the signal. Several examples are given and comparisons are made.

TABLE OF CONTENTS

	<u>Page</u>
ACKNOWLEDGMENT	1
ABSTRACT	2
LIST OF FIGURES	4
I. INTRODUCTION	6
II. THE MAXIMUM ENTROPY METHOD	8
III. EXTRAPOLATION OF BANDLIMITED SIGNALS	16
3.1 CONTINUOUS TIME SIGNALS	16
3.2 EXTENSION TO DISCRETE TIME SIGNALS	19
IV. EXTRAPOLATION OF DISCRETE TIME, BANDLIMITED SIGNALS	22
4.1 DEFINITIONS	22
4.2 PROPERTIES OF L	25
4.3 ITERATIVE EXTRAPOLATION	28
4.4 THE EXTRAPOLATION MATRIX	33
4.5 THE GENERALIZED INVERSE	36
4.6 DISCRETE PROLATE SPHEROIDAL WAVE FUNCTIONS AND SINGULAR VALUE EXPANSION	37
V. A CONJUGATE GRADIENT ALGORITHM FOR SIGNAL EXTRAPOLATION	43
VI. A MEAN SQUARE EXTRAPOLATING FILTER	48
VII. A RECURSIVE EXTRAPOLATION ALGORITHM	50
VIII. EXAMPLES, RESULTS AND COMPARISONS	54
IX. CONCLUSIONS	96
X. BIBLIOGRAPHY	101

# List of Figures

<u>Figure Number</u>	<u>Description</u>
1a.	Spectra of Signal, Clutter and Noise
1b.	Actual Clutter
1c.	Actual Signal
1d.	17 Samples of Clutter + Signal + Noise
1e.	Max Entropy Estimate Corresponding to 1d (8th order model)
1f.	256 pt. Power Spectrum of (1d) by FFT
1g.	Signal Extrapolated by Conjugate Gradient Algorithm
1h.	Signal Extrapolated by M.S. Extrapolation Filter
1i.	Clutter Extrapolated by Conjugate Gradient Algorithm
1j.	Clutter Extrapolated by M.S. Extrapolation Filter
1k.	Max Entropy Spectrum of (1g) (15th order model)
1l.	Max Entropy Spectrum of (1h) (15th order model)
1m.	Max Entropy Spectrum of (1i) (15th order model)
1n.	Max Entropy Spectrum of (1j) (15th order model)
2a.	Spectra of Signal and Clutter
2b.	256 pt. Power Spectrum of Observations by FFT
2c.	Max Entropy Spectrum of Observations by an 8th Order Model
2d.	Max Entropy Spectrum of Signal Extrapolated to 125 pts. (Ex 2)
2e.	Max Entropy Spectrum of Clutter Extrapolated to 125 pts. (Ex 2)
3,4,5,6,7,8,9,10 a	Original Signal
3,4,5,6,7,8,9,10 b	Given Observations (17 Samples)
3,4,5,6,7,8,9,10 c	Signal Extrapolated by Papoulis' Iterative Algorithm

<u>Figure</u>	<u>Description</u>
3,4,5,6,7,8,9,10 d	Signal Extrapolated Via Matrix $E_c$
3,4,5,6,7,8,9,10 e	Signal Extrapolated After Adding a Stabilizing Diagonal Term to Matrix $E_c$
3,4,5,6,7,8,9,10 f	Signal Extrapolated by Conjugate Gradient Algorithm
9,10 g	Signal Extrapolated by M.S. Extrapolation Filter
11	$\Lambda(f)$ vs. $f$

## I. INTRODUCTION

Spectral estimation refers to the problem of estimating the spectral density function of a stationary random signal which is observable only over a finite duration. For a deterministic signal it implies estimation of its magnitude spectrum. In either case, if the signal were known over the infinite interval, the Fourier transform of the signal or its autocorrelation would immediately yield the spectrum. Thus, any estimated spectrum is equivalent to specifying the signal or its autocorrelation outside the observation interval--i.e., its extrapolation.

In this report we consider several algorithms for extrapolation and spectral estimation of discrete time signals. First, we briefly review the maximum entropy (ME) or the linear predictive autoregressive (AR) method, and some iterative and matrix inverse based extrapolation algorithms developed recently by Papoulis [8], Sabri and Steenaart [10], and Cadzow [11].

The new results presented here are as follows.

- 1) Papoulis' iterative algorithm applied in discrete time domain converges to an extrapolated signal which is a minimum norm least squares type solution. It is seen to be a special case of a one step gradient algorithm, and has linear convergence. The convergence of this can be improved by suitably modifying it to a steepest descent algorithm.
- 2) Sabri and Steenaart [10] have reformulated Papoulis' iterative algorithm in terms of an extrapolation matrix operator  $E_{\infty}$  which yields the extrapolated signal when it operates on the given time truncated signal. It is proven that the infinite operator,  $E_{\infty}$  does not exist, but its finite truncation  $E_N$  exists, but it is ill-conditioned.

- 3) It is known that a continuous (time) band limited signal given over a finite observation interval can be extrapolated exactly outside this interval by means of Prolate Spheroidal Wave Functions (PSWF). We show that for the discrete time case a similar expansion arises when we consider the minimum norm least squares extrapolation.

Then we present three other algorithms which are as follows:

- 4) Conjugate Gradient Iterative Extrapolation
- 5) Minimum Mean Square Extrapolation Filter
- 6) Recursive Least Squares Extrapolation Filter

The conjugate gradient method is an iterative algorithm which yields a psuedo inverse extrapolation operator. Compared to the earlier iterative methods [8-11], this algorithm converges quite rapidly. The minimum mean square extrapolation algorithm is designed for applications where the observed band limited signal is contaminated by wideband white noise. It yields a simple, Wiener filter type, extrapolation operator which requires inversion of a matrix whose size is equal to the number of samples in the observed signal. No iterations are required here and the algorithm is shown to reduce to the matrix inverse algorithm of Cadzow [11] as the additive noise power goes to zero. Finally, the recursive least squares algorithm is a Kalman filter based method where the extrapolated signal estimate is updated recursively as a new observation sample arrives. The latter two methods are applicable in the presence of noise and yield stable results. Finally these algorithms are shown to be applicable to problems where one needs to discriminate as well as extrapolate an interfering signal and a desired signal.

Several examples are considered to compare the various algorithms.

## II. THE MAXIMUM ENTROPY METHOD [2-7]

Let  $\{u_i\}$  denote a real, zero mean, stationary, Gaussian random process whose covariance function is defined as

$$r_m = E[u_i u_{i+m}]. \quad (1)$$

We know  $r_m$  only on a finite window  $W$  defined as

$$W = \{-p \leq m \leq p\}. \quad (2)$$

The maximum entropy method extrapolates  $r_m$  outside  $W$  by maximizing the entropy

$$\epsilon \triangleq \int_{-1/2}^{1/2} \ln S(f) df \quad (3)$$

under the constraint

$$r_m = \int_{-1/2}^{1/2} S(f) e^{j2\pi mf} df, \quad m \in W. \quad (4)$$

The solution gives the maximum entropy spectrum as

$$S(f) = \frac{\beta^2}{\left[ \sum_{m \in W} a_m e^{-j2\pi fm} \right]^2}, \quad a_m = a_{-m}. \quad (5)$$

This could be written as

$$S(f) = \frac{\beta^2}{\left| \sum_{m=0}^p \alpha_m e^{-j2\pi fm} \right|^2} \quad (6)$$

where the  $a_m$  and  $\alpha_m$  are related by

$$a_{-m} = a_m = \sum_{k=\max[0, -m]}^{\min(p-m, m)} \alpha_{k+m} \alpha_k. \quad (7)$$

The coefficients  $\{\alpha_m\}$  are determined by solving (4) and (6) which is equivalent to solving (9) below. Alternatively, the  $\{u_i\}$  could be characterized as an AR process

$$u_k = \sum_{n=1}^p \alpha_n u_{k-n} + \epsilon_k \quad (8)$$

where  $\beta^2 = E[\varepsilon_k^2]$ . By writing the Yule-Walker equations for (8) it is a simple matter to show that  $\{\alpha_m\}$  are obtained by solving a set of simultaneous linear equations\*

$$R\alpha = -\beta^2 \underline{1}, \quad \beta^2 = 1/(R^{-1})_{1,1} \quad (9)$$

$$\underline{1} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \underline{\alpha} = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_p \end{bmatrix}, \quad \alpha_0 \triangleq -1$$

where  $\underline{\alpha}$  and  $\underline{1}$  are  $(p+1) \times 1$  vectors and  $R$  is a  $(p+1) \times (p+1)$  covariance matrix with entries corresponding to covariances on the window  $W$ , i.e.,

$$R = \begin{bmatrix} r_0 & r_1 & \cdot & \cdot & \cdot & r_p \\ r_1 & r_0 & r_1 & \cdot & \cdot & \cdot \\ r_2 & r_1 & r_0 & r_1 & \cdot & \cdot \\ \cdot & \cdot & r_1 & r_0 & r_1 & \cdot \\ \cdot & \cdot & \cdot & r_1 & r_0 & r_1 \\ r_p & \cdot & \cdot & \cdot & r_1 & r_0 \end{bmatrix}$$

\*For a positive definite matrix  $R$ ,  $\{\alpha_m\}$  are guaranteed to be such that  $S(f)$  is positive and (8) is asymptotically, a stationary random process.

Strictly speaking, the covariance values  $\{r_m, m \in W\}$  should be known exactly. In practice, one only knows data values on a finite window. Then, the covariances could be estimated as\*

$$r_m \approx \frac{1}{M} \sum_{k=1}^{M-|m|} u_k u_{k+|m|}, \quad m \in W \quad (10)$$

where  $M$  is the size of the data window. For large  $M \gg p$ , reasonable estimates of  $\{r_m\}$  could be expected.

Note that this method does not require  $\{u_i\}$  to be bandlimited (with respect to the Nyquist rate). Also, the spectral density function is, in view of (5) and (6), an all pole model. Thus, if the given observations were of the form

$$y_k = u_k + n_k \quad (11)$$

where  $n_k$  is a white noise process or another signal (e.g. clutter noise which could be modeled by an AR process), the spectrum of  $\{y_k\}$  would not be an all pole model and may have to be approximated by a very high order all pole model.

Example 1(a): Although there are many examples where the maximum entropy method could be applied successfully [3,5,7] we consider a case where it does not. We assume the observations to be given by

$$y_k = s_k + c_k + n_k \quad (12)$$

where  $s_k$  represents a bandlimited signal whose spectrum lies in the interval  $[f_2, f_3]$  and  $[-f_2, -f_3]$  and  $c_k$  is an interference signal bandlimited in the interval  $[-f_1, f_1]$  and  $n_k$  is a white noise process.

---

\*In estimating  $r_m$ , the divisor of  $M$ , rather than  $M-|m|$  is recommended. Although this results in a biased estimate of  $r_m$ , it yields a positive definite sequence  $\{r_m\}$  so that  $R$  is positive definite and the resulting spectra is positive. See Parzen [15] for details.

Figure 1(a) shows the spectra of the various signals. Figure 1(b) and 1(c) show the interference signal  $c_k$  and the actual signal  $s_k$ , modeled as

$$s_k = 1.69 \sin(.39\pi k).$$

Figure 1(d) shows the 17 samples of the available observations.

The signal to interference signal (to be called clutter) ratio, which is defined as

$$SCR^* = 20 \log_{10} \frac{\text{Peak to Peak Value of Signal}}{\text{r.m.s. value of clutter}}, \quad (13)$$

is -4.1 dB and the signal to noise ratio, SNR, defined similarly is 19 dB.

Figure 1(e) shows the maximum entropy spectrum estimate. A peak is expected at the position marked by the arrow. At this point the signal estimate is 30 dB below the clutter peak and is indistinguishable from the interfering signal. Figure 1(f) shows the spectrum estimated by directly evaluating the Fourier spectrum (i.e. the periodogram) as

$$S(f) = \frac{1}{17} \left| \sum_{k=-8}^8 y_k \exp(-j2\pi f k) \right|^2, \quad -\frac{1}{2} \leq f \leq \frac{1}{2} \quad (14)$$

Equation (14) can be evaluated approximately by discretizing the variable  $f$  and using a fast Fourier transform algorithm. The spectrum of Figure 1(f) is the result of a 256 point FFT. We note that both of these estimates are unsatisfactory.\* We will see that the new algorithms introduced here improve the estimated signal spectrum considerably.

---

\*Note that for random signals, the periodogram is an inconsistent estimate. Windowing techniques may be used to improve the spectrum estimate in the sense that it would be a consistent estimate of a smoothed version of the original spectrum. In this example, windowing did not improve the situation in so far as the signal was concerned.

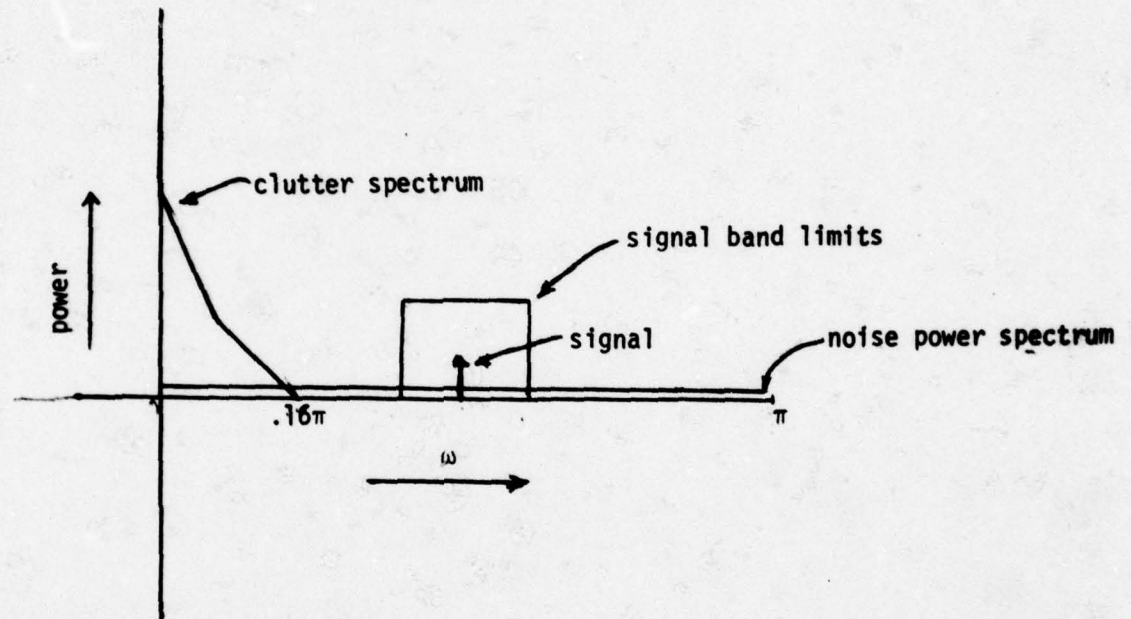


Fig. 1a: Spectra of Signal, Clutter and Noise

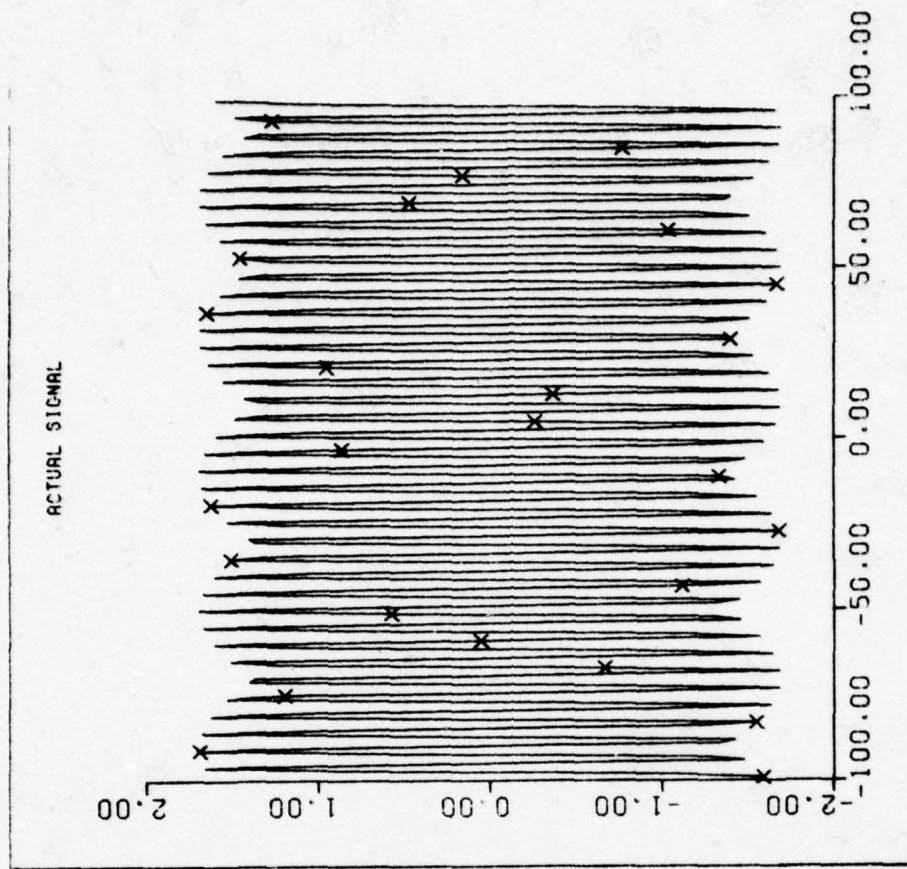


Fig. 1c: Actual Signal

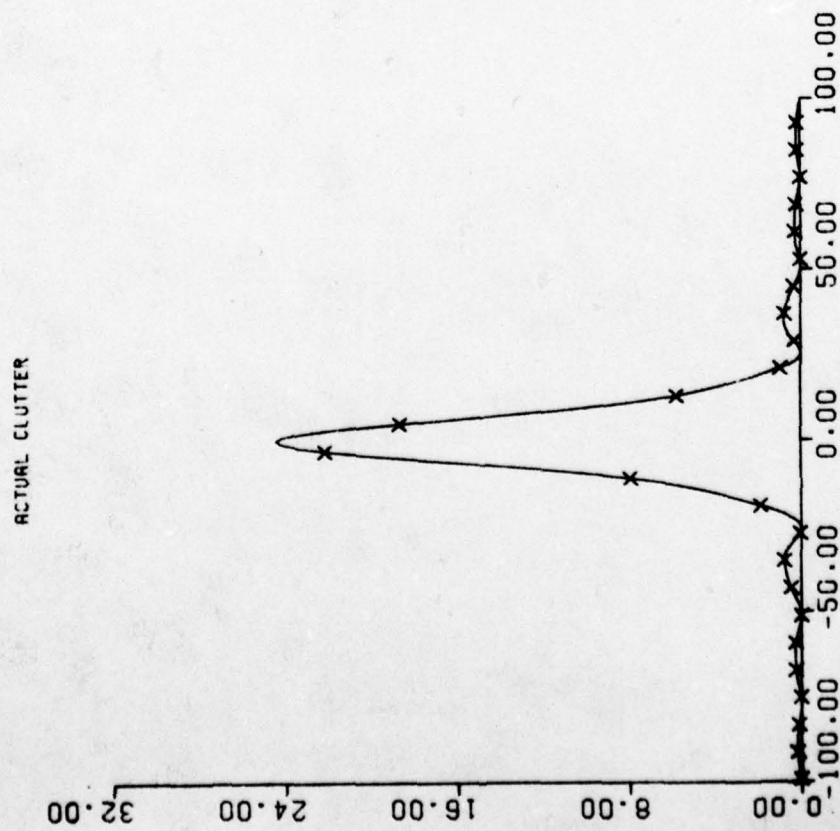


Fig. 1b: Actual Clutter

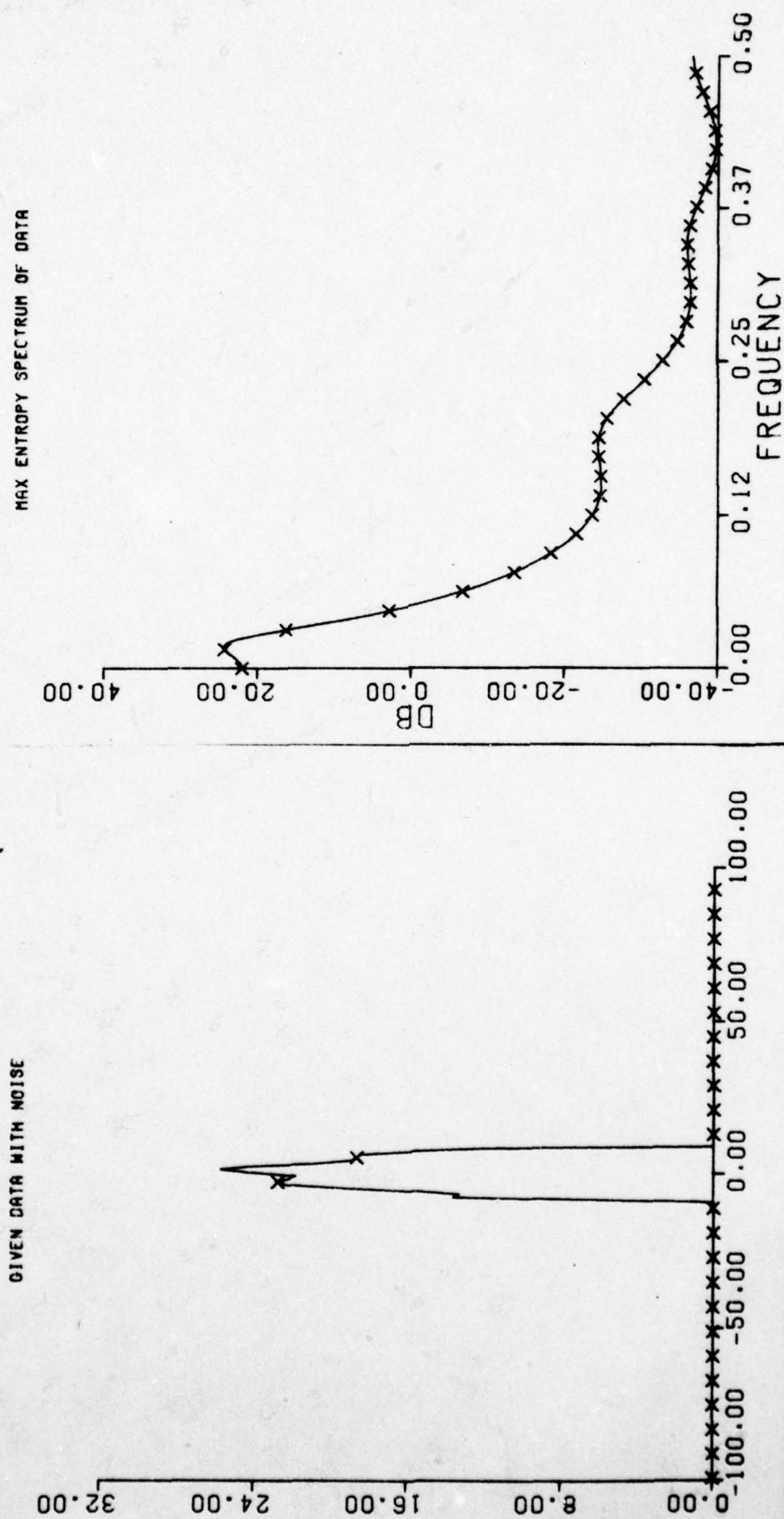
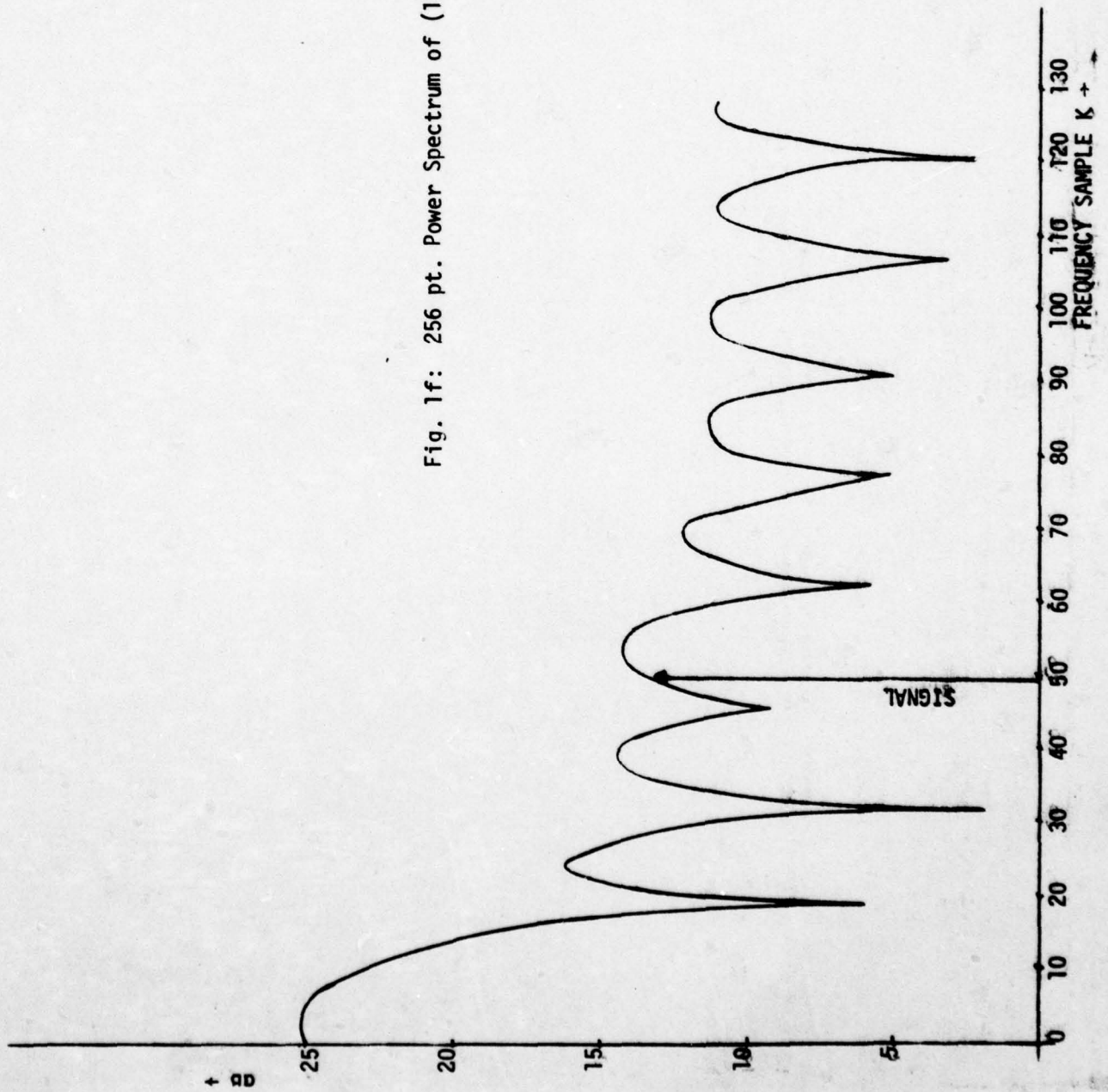


Fig. 1e: Max Entropy Estimate Corresponding to 1d  
(8th order model)

Fig. 1d: 17 Samples of Clutter + Signal + Noise

FIG. 1f

Fig. 1f: 256 pt. Power Spectrum of (1d) by FFT



### III. EXTRAPOLATION OF BANDLIMITED SIGNALS

#### 3.1 Continuous Time Signals

Suppose we have a continuous, bandlimited function  $f(t)$  so that its Fourier transform satisfies

$$F(\omega) = 0 \quad \text{for} \quad |\omega| > \sigma. \quad (15)$$

Let  $g_0(t)$  be a time limited segment of  $f(t)$  which is available as a noise-free observation, viz.,

$$g_0(t) = \begin{cases} f(t) & , \quad |t| \leq T \\ 0 & , \quad |t| > T \end{cases} \quad (16)$$

The problem is to extrapolate  $g_0(t)$  outside the interval  $[-T, T]$ . This is the classical problem of extrapolation of analytic functions. The existence of a unique solution can be established by observing that bandlimitedness of  $f(t)$  implies it is analytic. This means all its derivatives exist and are bounded so that from the Taylor series expansion

$$f(T+\Delta) = f(T) + \Delta f'(T) + \frac{\Delta^2}{2} f''(T) + \dots \quad (17)$$

one can evaluate  $f(t)$  outside  $[-T, T]$ . In practice, (17) is not very useful, because, not only does the series have to be truncated, but also, the evaluation of various derivatives is a noise sensitive process. An alternative algorithm suggested by Slepian, et al. [14], uses a series expansion

$$f(t) = \sum_{n=0}^{\infty} a_n \phi_n(t, T\sigma) \quad (18)$$

where  $\{\phi_n(t, T\sigma)\}$  is a special set of complete orthonormal bandlimited functions, called the prolate-spheroidal wave functions (PSWF), which are defined for all  $t$ . The coefficients  $\{a_n\}$  can be evaluated as projections of the known function  $g_0(t)$  on the basis functions  $\{\phi_n\}$ . Once  $\{a_n\}$  are known, the right side of (18), considered valid for all  $t$ , gives the extrapolated signal. In practice, this method also suffers from noise limitations and errors due to truncation of the series. Moreover it is extremely difficult to accurately generate the basis PSW functions so that extrapolation in a practical situation is quite hopeless. For a simple example, see Frieden [19].

Recently, Papoulis [8] has introduced an iterative scheme that appears to do better than the PSWF expansion method. The algorithm has the following steps. The first step is to compute the Fourier transform of  $g_0(t)$  as  $G_0(\omega)$  and define

$$F_1(\omega) = \begin{cases} G_0(\omega), & |\omega| < \sigma \\ 0 & , |\omega| > \sigma \end{cases} \quad (19)$$

Compute  $f_1(t)$ , the Fourier inverse of  $F_1(\omega)$  and let

$$g_1(t) = \begin{cases} f(t), & |t| < T \\ f_1(t), & |t| > T \end{cases} \quad (20)$$

Then compute  $G_1(\omega) = F[g_1(t)]$ .

This is the first step of the iteration. At the  $n^{\text{th}}$  step form the function

$$F_n(\omega) = \begin{cases} G_{n-1}(\omega), & |\omega| < \sigma \\ 0 & , |\omega| > \sigma \end{cases} \quad (21)$$

Find  $F^{-1}[F_n(\omega)]$  and form

$$g_n(t) = \begin{cases} f(t), & |t| < T \\ f_n(t), & |t| > T. \end{cases} \quad (22)$$

Papoulis has theoretically shown that  $f_n(t)$  converges to  $f(t)$  as  $n \rightarrow \infty$ . If we define a band-limiting operator as

$$Bf(t) = f(t) \otimes ((\sin \omega t)/\pi t) \quad (23)$$

where  $(\sin \omega t)/\pi t$  represents the impulse response of a low pass filter, and we define a time-limiting operator as

$$Df(t) = \begin{cases} f(t), & |t| \leq T \\ 0, & |t| > T \end{cases} \quad (24)$$

$$\bar{D} = I - D$$

$$\text{where } I = \text{identity operator} \quad (25)$$

then the foregoing algorithm can be written as [10]

$$\begin{aligned} g_n(t) &= g_0(t) + Hg_{n-1}(t) \\ f_n(t) &= f_1(t) + Gf_{n-1}(t) \end{aligned} \quad (26)$$

$$H = \bar{D}B, \quad G = B\bar{D}$$

or

$$\begin{aligned} f_n(t) &= \left[ \sum_{k=0}^{n-1} G^k \right] f_1(t), \quad G^0 = I \\ g_n(t) &= \left[ \sum_{k=0}^n H^k \right] g_0(t). \end{aligned} \quad (27)$$

Either  $g_n(t)$  or  $f_n(t)$  may be considered as the extrapolated signal.

### 3.2 Extension to Discrete Time Signals

Sabri and Steenaart [10] have suggested a discrete version of this algorithm, as follows.

Let  $y(k)$  be a discrete, bandlimited signal so that its Fourier transform (i.e., Z transform evaluated on the unit circle) defined as

$$Y(f) = \sum_{k=-\infty}^{\infty} y(k) \exp(-j2\pi f k), \quad -\frac{1}{2} \leq f \leq \frac{1}{2} \quad (28)$$

satisfies the relation\*

$$Y(f) = 0, \quad \frac{1}{2} > |f| > \sigma \quad (29)$$

We are given a set of time limited, noise free observations

$$g_0(k) = \begin{cases} y(k), & -M \leq k \leq M \\ 0, & \text{otherwise} \end{cases}$$

Given  $\{g_0(k)\}$ , the problem is to find an estimate of  $y(k)$  outside the interval  $[-M, M]$ . Following section 3.1, we define infinite vectors

$$\begin{aligned} y &= [\dots y(-k) \dots y(-1), y(0), y(1), \dots, y(k) \dots]^T \\ g_n &= [\dots g_n(-k) \dots g_n(-1), g_n(0), g_n(1), \dots, g_n(k) \dots]^T \end{aligned} \quad (30)$$

where  $g_0 \triangleq [0, 0, \dots, 0, g_0(-M), g_0(-M+1), \dots, g_0(-1), g_0(0), g_0(1), \dots, g_0(M), 0, 0, \dots, 0]^T$ .

We also define a band-limiting operator  $L$ , and a time-limiting operator  $W$ , as infinite matrices

\*This implies  $y(k)$  has been oversampled with respect to its Nyquist rate. This occurs quite often when a system observes signals over a wide bandwidth.

$$L = \{l_{i,j}\}, \quad l_{i,j} = \frac{\sin 2\pi(i-j)\sigma}{\pi(i-j)}, \quad i, j = 0, \pm 1, \pm 2, \dots \quad (31)$$

$$W = \{w_{i,j}\}, \quad w_{i,j} = \begin{cases} 1, & i=j, \quad -M \leq i, j \leq M \\ 0 & , \text{ otherwise} \end{cases}$$

Also, let

$$\bar{W} = I - W \quad (32)$$

Then, one obtains from (27)

$$f_{n+1} = \left[ \sum_{k=0}^n G^k \right] f_1 \quad (33)$$

$$G = L\bar{W}$$

$$f_1 = Lg_0$$

Defining

$$E_n \triangleq \sum_{k=0}^n G^k \quad (34)$$

in the limit as  $n \rightarrow \infty$ , we get

$$f_\infty = E_\infty f_1 \quad (35)$$

where

$$E_\infty = \sum_{k=0}^{\infty} G^k = (I-G)^{-1}. \quad (36)$$

The matrix  $E_\infty$  has been called the extrapolation matrix and it exists if and only if no roots of  $G$  equal unity. In iterative form, the algorithm becomes

$$f_{n+1} = f_1 + Gf_n \quad (37)$$

In practice, the infinite matrix operator  $G$  is replaced by a finite matrix, of size, say,  $N \times N$  where  $N \gg (2M+1)$ . Later in this section we will show that  $(I-G)$  is singular for  $N = \infty$ , but its finite approximation ( $N < \infty$ ) is non singular.

### Existence and Convergence

Following Papoulis, it can be shown that the above iterative algorithm satisfies the inequality

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} |Y(f) - G_{n-1}(f)|^2 df > \int_{-\frac{1}{2}}^{\frac{1}{2}} |Y(f) - G_n(f)|^2 df \quad (38)$$

which says that the mean square error is reduced at each step. However, the extrapolated signal need not converge to the original signal  $y(k)$  because the time limited discrete signal does not have the analyticity property that the continuous signal has. Indeed, as we show in the next section, the above (discrete extrapolation) algorithm converges to a least squares, minimum norm solution associated with the solution of the equation

$$WLy = g_0$$

In terms of computational complexity, the iterative algorithm requires about  $4nN \log_2 N$  real operations (one operation = one multiplication and one addition), where  $N$  is the size of the extrapolated vector (and is much larger than  $M$ ) and  $n$  is the number of iterations.

If the extrapolation matrix  $E_m$  is used, then once it has been computed, it requires  $\frac{1}{2}(2M+1)(N-2M-1)$  operations to evaluate the extrapolated signal. However a large ( $N \times N$ ) matrix which is ill conditioned, has to be inverted. (see next section)

#### IV. EXTRAPOLATION OF DISCRETE TIME, BANDLIMITED SIGNALS

Before proceeding to prove several results related to extrapolation of discrete signals we first consider several definitions. Let  $A$  be an arbitrary  $m \times n$  matrix and consider the system of equations

$$Ay = z \quad (39)$$

where  $y$  and  $z$  are  $n \times 1$  and  $m \times 1$  vectors respectively.

##### 4.1 Definitions

##### Definition 1: Least Squares Solution

A least squares solution of (39), denoted by  $\hat{y}$ , is such that

$$\|z - A\hat{y}\|^2 = (z - A\hat{y})^T (z - A\hat{y}) \quad (40)$$

is minimum. This solution must therefore satisfy the equation

$$A^T A \hat{y} = A^T z \quad (41)$$

If  $A^T A$  is nonsingular (i.e.  $n \leq m$  and rank of  $A$  is  $n$ ) then

$$\hat{y} = (A^T A)^{-1} A^T z \quad (42)$$

is the least squares solution and is unique. If  $m = n$  and  $A$  is nonsingular, then

$$\hat{y} = y = A^{-1} z \quad (43)$$

If  $m < n$ , then  $A^T A$  is necessarily singular and has rank at most  $m$ . Then (41) does not have a unique solution.

Definition 2: Minimum Norm Least Squares Solution

Let  $y^+$  denote this solution. Then  $y^+$  must be that solution of (41) which has the minimum norm  $\|\hat{y}\|^2$ . Thus

$$y^+ = \min_{\hat{y}} \{ \|\hat{y}\|^2; A^T A \hat{y} = A^T z \} \quad (44)$$

Clearly, if rank  $A^T A$  is  $n$  then  $y^+ = \hat{y}$ . The minimum norm is simply a constraint that makes the least squares solution unique for an arbitrary  $A$ .

Definition 3: Pseudo Inverse

We call  $A^+$  the pseudo inverse of  $A$  [18], if for every equation (39), the associated minimum norm least squares solution is given by

$$y^+ = A^+ z \quad (45)$$

This pseudo inverse, also called the generalized inverse, satisfies the relations

$$\begin{aligned} AA^+ &= (AA^+)^T \\ A^+A &= (A^+A)^T \\ AA^+A &= A \\ A^+AA^+ &= A^+ \end{aligned} \quad (46)$$

When rank  $A = n$ ,

$$A^+ = (A^T A)^{-1} A^T \quad (47)$$

If rank  $A = m$ , then

$$A^+ = A^T (A A^T)^{-1} \quad (48)$$

Definition 4: Singular Value Expansion [21]

In general an explicit expression for  $A^+$ , of the type of (47) or (48) is not available. However,  $A^+$  can be expressed as an expansion. Consider the eigenvalue problems

$$A^T A \phi_k = \lambda_k \phi_k \quad (49)$$

$$A A^T \psi_k = \lambda_k \psi_k$$

where  $k = 1, 2, \dots, p$  and  $p$  is the rank of  $A$ . The vectors  $\phi_k$  and  $\psi_k$  are of sizes  $n \times 1$  and  $m \times 1$  respectively. Since  $A^T A$  and  $A A^T$  are non-negative matrices, these eigen-vectors exist and can be orthonormalized so that

$$\phi_k^T \phi_\ell = \delta_{k,\ell} \quad (50)$$

$$\psi_k^T \psi_\ell = \delta_{k,\ell}$$

From this, one can express the rectangular matrix  $A$  by the expansion, called the singular value expansion, as

$$A = \sum_{k=1}^P \lambda_k \psi_k \phi_k^T \quad (51)$$

where  $\lambda_k > 0$ , are called the singular values of A.

The pseudo inverse  $A^+$  can now be written as

$$A^+ = \sum_{k=1}^P \frac{1}{\lambda_k} \phi_k \psi_k^T \quad (52)$$

#### 4.2 Properties of L

Now we consider some useful properties of the low pass filter operator defined in (31).

Property I: L is a symmetric operator, i.e.,  $L = L^T$ . This follows from its definition.

Property II: The Fourier spectrum of L is given by

$$\ell(f) = \begin{cases} 1, & 0 < |f| < \sigma \\ 0, & \text{otherwise,} \end{cases} \quad (53)$$

where  $-\frac{1}{2} \leq f \leq \frac{1}{2}$ . This is obvious since L is the Toeplitz matrix formed by the Fourier inverse of  $\ell(f)$ , the lowpass filter transfer function.

Property III: Let S be a  $(2M+1) \times \infty$  matrix operator whose elements are

$$S_{i,j} = \begin{cases} 1, & i=j=0 \pm 1, \pm 2, \dots, \pm M \\ 0, & \text{otherwise} \end{cases} \quad (54)$$

Basically S selects  $(2M+1)$  elements from an infinite vector. Consider the  $(2M+1) \times (2M+1)$  matrix

$$\hat{L} = SLS^T \quad (55)$$

Then

$$\hat{\ell}_{i,j} = \hat{\ell}_{i-j} = \frac{\sin 2\pi\sigma(i-j)}{\pi(i-j)}; \quad -M \leq i,j \leq M \quad (56)$$

Also

$$S^T S = W \quad (57)$$

where W is defined in (31).

Property IV: The operator L is idempotent, i.e.,

$$L^2 = L \quad (58)$$

This is obvious because ideal low pass filtering a signal once is the same as doing it twice, i.e.,

$$Ly = L(Ly)$$

Note this implies the spectrum (or eigenvalues) of L must be composed of zeros and ones only [see (53)].

Property V: For every  $M < \infty$ ,  $\hat{L}$  is positive definite. This follows by noting

$$x^T \hat{L} x^* = \sum_m \sum_n x_m \hat{\ell}_{m-n} x_n^* = \sum_m \sum_n \left( \int_{-\sigma}^{\sigma} e^{j2\pi(m-n)f} df \right) x_m x_n^*$$

$$\begin{aligned}
 &= \int_{-\sigma}^{\sigma} \left| \sum_{m=-M}^M x_m e^{j2\pi m f} \right|^2 df \\
 &\triangleq \int_{-\sigma}^{\sigma} |x_M(f)|^2 df, \quad x_M(f) \triangleq \sum_{m=-M}^M x_m e^{j2\pi m f} \quad (59) \\
 &> 0, \quad \text{if } M < \infty
 \end{aligned}$$

If  $M = \infty$ , then  $x_m = e^{-j2\pi m \xi}$  gives  $x_M(f) = \delta(f - \xi)$  so that

$$x^T \hat{L} x^* = x^T L x^* = \begin{cases} 1, & |\xi| < \sigma \\ 0, & |\xi| > \sigma \end{cases} \quad (60)$$

and is not positive definite. Thus, all the eigenvalues of  $\hat{L}$  are positive

$$\lambda(\hat{L}) > 0, \quad M < \infty \quad (61)$$

Property VI: Let  $\lambda_{\max}(\hat{L})$  denote the largest eigenvalue of  $\hat{L}$ . Then

$$\begin{aligned}
 \lambda_{\max}(\hat{L}) &< 1, \quad M < \infty \\
 &= 1, \quad M = \infty
 \end{aligned} \quad (62)$$

Thus, for any finite  $M$ , the eigenvalues of  $\hat{L}$  are bounded in the interval  $(0, 1)$  i.e.,

$$0 < \lambda(\hat{L}) < 1, \quad M < \infty \quad (63)$$

To prove (62) we note

$$\lambda_{\max}(\hat{L}) \triangleq \max_{\{x\}} \frac{x^T \hat{L} x^*}{x^T x^*}$$

From (59) we can write

$$\lambda_{\max}(\hat{L}) = \max_{\{x_m\}} \frac{\int_{-\sigma}^{\sigma} |x_m(f)|^2 df}{\int_{-1/2}^{1/2} |x_M(f)|^2 df} \quad (64)$$

Since, for  $M < \infty$ ,  $x_M(f)$  is the Fourier spectrum of a time limited signal, it cannot be zero on any finite interval. Hence

$$\int_{-\sigma}^{\sigma} |x_M(f)|^2 df < \int_{-1/2}^{1/2} |x_m(f)|^2 df, \quad \forall M < \infty, \sigma < \frac{1}{2}$$

$$\text{or } \sum_m \sum_n x_m \hat{e}_{m-n} x_n^* < \sum_{m=-M}^M |x_m|^2, \quad \forall M < \infty.$$

When  $M = \infty$ , one could maximize (64) by choosing a bandlimited signal so that the above inequality will become an equality. This proves (62).

#### 4.3 Iterative Extrapolation

With the above definitions and properties we are now ready to prove the following results. Let  $y(k)$ ,  $k=0, \pm 1, \dots$  be a discrete time bandlimited signal as defined in (28) and (29).

Furthermore, let this signal be observed without any noise over a finite interval and define

$$z(k) = y(k), \quad -M \leq k \leq M \quad (65)$$

If  $z$  denotes a  $(2M+1) \times 1$  vector and  $y$  is the infinite vector of  $\{y(k)\}$ , then

$$z = Sy$$

Since  $y$  is bandlimited, it must satisfy

$$Ly = y$$

so that we can write

$$z = SLy \quad (66)$$

Theorem 1: Minimum Norm Least Squares Extrapolation Theorem

The iterative solution [see (32) to (37) and (57)]

$$\begin{aligned} f_{q+1} &= f_1 + Gf_q, \quad q = 1, 2, \dots \\ G &= L(I - S^T S) = L(I - W) \\ f_1 &= Lg_0 = LS^T z \end{aligned} \quad (67)$$

converges to the minimum norm least squares solution  $y^+$  of (66). Moreover, (67) is a special case of a gradient algorithm associated with the minimum norm least squares optimization problem.

Proof: An iterative gradient algorithm associated with the minimum norm least squares solution of the general equation (39) is

$$y_{q+1} = y_q + \frac{1}{\sigma} A^T (z - Ay_q) \quad (68)$$

$$= \frac{1}{\sigma} A^T z + (I - \frac{1}{\sigma} A^T A) y_q \quad (69)$$

It is known that  $y_q$  converges to  $y^+$  as  $q \rightarrow \infty$ , under the following conditions [12]:

$$a) \quad 0 < \frac{1}{\sigma} < \frac{2}{\lambda_{\max}(A^T A)} \quad (70)$$

b) The initial guess  $y_0$  must lie in the range space of  $A^T A$  e.g.,  $y_0 = 0$ .

From (66), letting

$$A = SL \quad (71)$$

we get

$$A^T A = L^T S^T SL = L^T L = I \quad (72)$$

$$A^T z = L^T S^T z = L^T z = f_1 \quad (73)$$

Hence (69) becomes

$$y_{q+1} = \frac{1}{\sigma} f_1 + y_q - \frac{1}{\sigma} L^T L y_q \quad (74)$$

Now letting

$$y_0 = 0$$

and noting that  $f_1$  is bandlimited i.e.,

$$L f_1 = f_1 \quad (75)$$

it is easily verified by induction that  $\{y_q\}$  is a bandlimited sequence i.e.,

$$Ly_q = y_q, \quad q=1,2,3\dots \quad (76)$$

Using this in (74) we get

$$\begin{aligned} y_{q+1} &= \frac{1}{\sigma} f_1 + y_q - \frac{1}{\sigma} LWy_q \\ &= \frac{1}{\sigma} f_1 + (I - \frac{1}{\sigma} LW)y_q \end{aligned} \quad (77)$$

For  $\sigma = 1$ , (77) becomes the same as (67). Now it remains to show that this algorithm converges for  $\sigma = 1$ . From (70) and (72) this requires us to find the largest eigenvalue of  $LWL$ . Now

$$\begin{aligned} \lambda_{\max}(LWL) &= \lambda_{\max}(A^T A) = \lambda_{\max}(AA^T) \\ &= \lambda_{\max}(SLS^T) \\ &= \lambda_{\max}(\hat{L}) \\ &< 1, \quad \forall M < \infty. \end{aligned}$$

where we have used Property VI. Therefore, convergence of (77) is achieved whenever

$$0 < \frac{1}{\sigma} \leq 2 < \frac{2}{\lambda_{\max}(\hat{L})} \quad (78)$$

Hence for  $\sigma = 1$ , (77) converges. This completes the proof of Theorem 1.

An interesting question raised by the foregoing result is "What is the optimal value of  $\sigma$ ?" In other words, we want to find the "steepest descent" for the gradient algorithm. Defining the error vector at iteration step  $q$  as

$$e_q = y - y_q \quad (79)$$

and noting that  $f_1$  can also be written as

$$f_1 = Lg_0 = LWy = LWLy \quad (80)$$

we obtain from (74)

$$e_{q+1} = (I - \frac{1}{\sigma} LWL)e_q \quad (81)$$

$$= (I - \frac{1}{\sigma} LWL)^{q+1} e_0, e_0 = y \quad (82)$$

This shows the convergence rate of this extrapolation algorithm is linear. Slow convergence of this algorithm has also been noted experimentally by us and by Papoulis [8] and Sabri et al [10]. Since this is a gradient algorithm, convergence can be improved by adjusting  $\sigma$  at every iteration. The optimal value is given by

$$\sigma_q^{-1} = \frac{h_q^T h_q}{h_q^T A^T A h_q} \quad A = SL, \quad (83)$$

where  $h_q$  is the gradient at step  $q$ , defined as

$$\begin{aligned} h_q &= A^T(z - Ay_q) \\ &= f_1 - LWy_q \end{aligned} \quad (84)$$

This requires additional computations at every iteration step. If a constant value of  $\sigma$  is desired, it is given by [24]

$$\sigma_{\text{opt}}^{-1} = 2/[\lambda_{\text{max}}(\text{LWL}) + \lambda_{\text{min}}(\text{LWL})]$$

Since

$$\lambda_{\text{min}}(\text{LWL}) = 0$$

$$\lambda_{\text{max}}(\text{LWL}) < 1$$

we can take

$$\sigma_{\text{opt}}^{-1} \approx 2 \leq 2/\lambda_{\text{max}} \quad (85)$$

From our foregoing analysis we conclude the following about Papoulis' iterative algorithm.

1. The solution converges to a minimum norm least squares solution. Note that continuous version of the algorithm converges to the original band limited signal  $y(t)$ , as proven by Papoulis [8]. This reinforces the fact that time limited discrete observations of a band limited signal need not give its exact extrapolation.
2. The algorithm is a gradient algorithm. Hence its convergence is linear and slow. It could be improved by going to the steepest descent algorithm.

#### 4.4 The Extrapolation Matrix

Now we consider the extrapolation matrix suggested by Sabri and Steenaart [10]. This is the doubly infinite matrix defined as [see eqns. (32) to (36)]

$$E_{\infty} = (I-G)^{-1} \quad (86)$$

$$G = L(I-W)$$

In a practical situation the matrix  $G$  is truncated to a finite, but large,  $N \times N$  matrix, say  $G_N$  defined as

$$G_N = L_N(I_N - W_N) \quad (87)$$

and the corresponding extrapolation matrix is

$$E_N = (I_N - G_N)^{-1}$$

We intend to show that for every finite  $N$ ,  $E_N$  exists. However,  $E_\infty$  does not exist. Thus as  $N$  goes to infinity the sequence  $\{E_N\}$  becomes an ill-conditioned set of matrices.

Lemma 1: For every finite  $N$ , the matrix  $P_N$  defined as

$$P_N = I - L_N + L_N W_N \quad (88)$$

is nonsingular. At  $N = \infty$ ,  $P_N$  is singular.

Proof: From Property V, the finite  $N \times N$  matrix  $L_N$  is positive definite.

Now consider the symmetric matrix

$$\begin{aligned} C_N &\triangleq P_N L_N \\ &= L_N - L_N^2 + L_N W_N L_N. \end{aligned} \quad (89)$$

Since all the eigenvalues of  $L_N$  lie in the interval  $(0,1)$  we have

$$\lambda_k(L_N^2) = \lambda_k^2(L_N) < \lambda_k(L_N) \quad (90)$$

Therefore, for any  $N \times 1$  vector  $x$ ,

$$x^T L_N x^* > x^T L_N^2 x^*$$

Also

$$x^T L_N W_N L_N x^* = \sum_{m=-M}^M \sum_{n=-M}^M x_m \ell_{m,n}^{(2)} x_n^* > 0$$

where  $\{\ell_{m,n}^{(2)}\}$  are the  $(2M+1) \times (2M+1)$  elements of  $L_N^2$ , which is positive definite. Clearly, then  $C_N$  is positive definite. Hence  $P_N = C_N L_N^{-1}$  exists and is nonsingular.

At  $N = \infty$ ,  $L_N$  is singular. Consider the eigenvectors of the equation

$$LWLx = \lambda x \quad (91a)$$

Since  $LWL$  is symmetric and is of rank  $(2M+1)$ , there exists an  $x$  such that

$$LWLx = 0, \quad x \neq 0 \quad (91b)$$

Moreover, for every such vector there exists a band-limited  $x$ ,

$$Lx = x$$

which is also a solution of (91a). Thus, for all such  $x$  we have

$$\begin{aligned}
 x^T P x &= x^T P L x = x^T L x - x^T L^2 x + x^T L W L x \\
 &= x^T L x - x^T L x + x^T L W L x \\
 &= x^T L W L x \\
 &= 0,
 \end{aligned}$$

Thus P is singular.

#### 4.5 The Generalized Inverse

Having noted that the foregoing approaches give a minimum norm least squares solution, one may attempt to find it directly. We recall that the given system of equations is

$$S L y = z \quad (92)$$

Defining  $A = S L$  to give

$$\begin{aligned}
 A A^T &= S L L^T S^T \\
 &= S L S^T \\
 &= \hat{L}
 \end{aligned} \quad (93)$$

We note that  $\hat{L}$  is positive definite (Property V). Hence from Definition 3 and Eqn. (48) we can write directly  $A^+ = A^T (A A^T)^{-1}$  which gives the extrapolation matrix

$$E_c = L^T S^T (S L S^T)^{-1} \quad (94)$$

and

$$y^+ = L^T S^T (S L S^T)^{-1} z \quad (95)$$

This form of solution was obtained for the extrapolation problem by Cadzow [11] by a different route. This method, we believe, is easier and more direct. We note that while the extrapolation matrix  $E_\infty$  did not exist, the extrapolation matrix  $E_c$  exists. Moreover,  $\hat{L}$  is only a  $(2M+1) \times (2M+1)$  Toeplitz matrix so that its dimensionality is much smaller. However, as  $M$  becomes large or for certain combinations of  $M$  and  $\sigma$ ,  $\hat{L}$  could be ill conditioned. Experimentally, the ill conditioning can be reduced by adding a small diagonal term to  $\hat{L}$ . This, however, will degrade the extrapolated estimate.

#### 4.6 Discrete Prolate Spheroidal Wave Functions and Singular Value Expansion

In an earlier section we had mentioned that a continuous band-limited signal could be extrapolated outside its observation interval, exactly, via the PSWF expansion. For the case of discrete signals, a similar expansion is possible for the minimum norm least squares extrapolated estimate. This is achieved via the singular value expansion described in Definition 4. Papoulis and Bertram [20] have introduced the discrete PSWF earlier for realization of digital filters whose impulse response is an all-zero model. However, they have not shown any extrapolation properties of these PSWFs. Here we introduce the discrete PSWFs which also extrapolate a discrete band limited signal (known over a finite duration) to an infinite, minimum norm least squares signal.

Following Definition 4 for  $A = SL$ , we consider the eigenvalue problems

$$A^T A \phi_k = L W L \phi_k = L S^T S L \phi_k = \lambda_k \phi_k, \quad -M \leq k \leq M \quad (96)$$

$$A A^T \psi_k = \hat{L} \psi_k = S L S^T \psi_k = \lambda_k \psi_k, \quad -M \leq k \leq M \quad (97)$$

where  $\lambda_k > 0$ , and  $\{\phi_k\}$  are  $\infty \times 1$  and  $\{\psi_k\}$  are  $(2M+1) \times 1$  orthogonal vectors i.e.,

$$\left. \begin{aligned} \phi_k^T \phi_l &= \delta_{k,l} \\ \psi_k^T \psi_l &= \delta_{k,l} \end{aligned} \right\} \quad (98)$$

From (96),  $\phi_k$  must be a band-limited signal satisfying the condition

$$L \phi_k = \phi_k \quad (99)$$

because  $\phi_k = (1/\lambda_k) L (W L \phi_k)$ , and  $Lz$  is band limited,  $\forall z$ . Now define

$$\xi_k = S L \phi_k = S \phi_k \quad (100)$$

Then (96) gives

$$L S^T \xi_k = \lambda_k \phi_k$$

or

$$S L S^T \xi_k = \lambda_k S \phi_k = \lambda_k \xi_k \quad (101)$$

This shows  $\xi_k$  satisfies the same equation as  $\psi_k$ . Hence  $\xi_k$  and  $\psi_k$  must be proportional i.e.

$$\xi_k = c_k \psi_k$$

Using the orthogonality condition (98) and the eqns. (99), (100) and (97) we find

$$c_k^2 = \lambda_k$$

or

$$c_k = \sqrt{\lambda_k} \quad (102)$$

giving

$$\boxed{\psi_k = \frac{1}{\sqrt{\lambda_k}} S \phi_k}, \quad \lambda_k > 0, \quad -M \leq k \leq M \quad (103)$$

Also from (96), this yields

$$\boxed{\phi_k = \frac{1}{\sqrt{\lambda_k}} L S^T \psi_k}, \quad -M \leq k \leq M \quad (104)$$

Equation (103) states that the  $(2M+1) \times 1$  vector  $\psi_k$  is simply obtained by selecting the  $(2M+1)$  elements  $\{\phi_k(m), -M \leq m \leq M\}$  of  $\phi_k$  and scaling them by  $\lambda_k^{-1/2}$ .

Eqn. (104) is remarkable in that the  $\infty \times 1$  vector  $\phi_k$  is obtained by simply low pass filtering the sequence  $\{\psi_k(m)\}$  and scaling the result by  $\lambda_k^{-1/2}$ . This means  $\phi_k$  is the extrapolation of  $\psi_k$ , obtained by simple low pass filtering and scaling. Also noteworthy is the fact that the sequence  $\{\phi_k(m), -\infty \leq m \leq \infty\}$  is orthogonal over the interval  $-M \leq m \leq M$  as well as over the infinite interval. This property is similar to that of the continuous PSWFs. The  $(2M+1) \times 1$  vectors  $\psi_k$  are easily obtained by solving the eigenvalue problem of (97) i.e.

$$\sum_{m=-M}^M \frac{\sin 2\pi\sigma(n-m)}{\pi(n-m)} \psi_k(m) = \lambda_k \psi_k(n), \quad -M \leq n, k \leq M \quad (105)$$

Once  $\{\psi_k\}$  are obtained,  $\{\phi_k\}$  are given by (104) i.e.,

$$\phi_k(n) = \frac{1}{\sqrt{\lambda_k}} \sum_{m=-M}^M \frac{\sin 2\pi\sigma(n-m)}{\pi(n-m)} \psi_k(m), \quad -\infty \leq n \leq \infty \quad (106)$$

combining (105) and (106) we find

$$\phi_k(n) = \begin{cases} \sqrt{\lambda_k} \psi_k(n), & -M \leq n \leq M, \\ \frac{1}{\sqrt{\lambda_k}} \sum_{m=-M}^M \frac{\sin 2\pi\sigma(n-m)}{\pi(n-m)} \psi_k(m), & \text{otherwise} \end{cases} \quad (107)$$

where  $-M \leq k \leq M$ . The extrapolated signal is obtained by applying (52) as

$$y^+ = A^+ z$$

or

$$y^+(m) = \sum_{k=-M}^M \frac{a_k}{\sqrt{\lambda_k}} \phi_k(m), \quad \forall m \quad (108)$$

where

$$a_k = \psi_k^T z = \sum_{m=-M}^M \psi_k(m) z_m = \sum_{m=-M}^M \psi_k(m) y(m) \quad (109)$$

The functions  $\{\psi_k(n)\}$  and  $\{\phi_k(n)\}$  have been obtained by Papoulis [20] and are called the discrete PSWF. However, their usefulness in obtaining minimum norm least squares extrapolation has not been noted earlier. Our extrapolation algorithm requires the following steps. First calculate  $(2M+1)$  orthonormalized vectors (each of size  $(2M+1)$ ) by solving (105). Next obtain the  $(2M+1)$  infinite size vectors  $\phi_k$  according to (104) by low pass filtering and scaling of  $\psi_k$ . Then (108) and (109) give the extrapolated estimate from the observations  $\{y(m), -M \leq m \leq M\}$ .

Properties of Discrete PSWFs: We now summarize properties of these functions.

1. Let  $L$  be a low pass filter operator and  $\hat{L}$  be a  $(2M+1) \times (2M+1)$  principal minor of it, i.e.

$$\hat{L}_{i,j} = L_{i,j} = L_{i-j} \quad i-j=0, \pm 1, \dots, \pm(2M)$$

Then the orthonormalized eigenvectors of  $\hat{L}$ , denoted by  $\psi_k$  form a complete orthonormal set of basis functions in  $(2M+1)$  dimensional vector space. Let

$$\begin{aligned} \hat{L}\psi_k &= \lambda_k \psi_k & -M \leq k \leq M \\ \lambda_k &> 0 \end{aligned} \tag{110}$$

2. The discrete PSWF are formed from  $\{\psi_k\}$  as  $\infty \times 1$  vectors  $\{\phi_k\}$  and are defined as

$$\phi_k = \frac{1}{\sqrt{\lambda_k}} LS\psi_k \quad -M \leq k \leq M \tag{111}$$

3. The PSWFs are orthonormal over the infinite interval, i.e.

$$\sum_{m=-\infty}^{\infty} \phi_k(m) \phi_l(m) = \delta_{k,l} \quad (112)$$

4. The PSWFs are orthogonal over the finite interval  $[-M, M]$ , i.e.,

$$\sum_{m=-M}^M \phi_k(m) \phi_l(m) = \lambda_k \delta_{k,l} \quad (113)$$

5. The Fourier transform  $\phi_k(f)$  of  $\{\phi_k(m)\}$  satisfies the eigenvalue integral equation

$$\int_{-\sigma}^{\sigma} \frac{\sin(2M+1)\pi(f-f')}{\sin\pi(f-f')} \phi_k(f') df' = \lambda_k \phi_k(f), \quad |f| < \sigma \quad (114)$$

The proof is obtained by Fourier transforming (111) and using (103).

6. The eigenvalues of  $\hat{L}$  lie in the interval  $(0, 1)$  i.e.,  $0 < \lambda_k(\hat{L}) < 1$ .

7. Let  $y(m)$  be a band limited signal whose spectrum lies in the interval  $(-\sigma, \sigma)$ . If  $y(k)$  is known over the interval  $[-M, M]$ , then its minimum norm least squares extrapolation estimate is given by

$$y^+(m) = \sum_{k=-M}^M \frac{a_k}{\lambda_k} \phi_k(m), \quad \forall m \quad (115)$$

$$a_k = \sum_{m=-M}^M y(m) \phi_k(m)$$

Note that this gives  $y^+(m) = y(m)$  for  $m \in [-M, M]$ .

So far we have considered only the case when there is no noise in the observed signal. In the next few sections, we consider other algorithms where additive noise or interfering signals are allowed in the observations. The algorithms reduce to minimum norm least squares solutions in the absence of any noise.

# V. A CONJUGATE GRADIENT ALGORITHM FOR SIGNAL EXTRAPOLATION

In this section we consider a slightly different but more general problem than that considered in the previous section. We are interested in extrapolation and discrimination of two interfering signals (see Example 1).

Let us assume that we are given  $m$  observations which may consist of signal, clutter and noise, where the signal and clutter are bandlimited in mutually exclusive bands (see Fig. 1(a)). The problem then is to obtain estimates of extrapolated signal and clutter outside their observation interval.

Let the Bandpass operator  $B$  operate on the signal in the signal band  $\pm[f_2, f_3]$  and the low pass operator  $L$  operate on the clutter in the clutter band  $\pm[0, f_1]$ .

We introduce operator notation:

- $s$ : original signal (infinite vector)
- $c$ : clutter (infinite vector)
- $n$ : noise ( $m \times 1$  vector)
- $y$ : observed samples ( $m \times 1$  vector)
- $B$ : bandpass operator (infinite matrix)
- $L$ : Low pass operator (infinite matrix)
- $S$ : Selection operator ( $m \times \infty$ )

The matrices  $B$  and  $L$  are Toeplitz matrices determined from the sequences  $\{b_k\}$ ,  $\{l_k\}$  respectively, as the Fourier inverses [see Fig. 11],

$$b_k = 2 \int_{f_2}^{f_3} \cos(2\pi f k) df = \frac{\sin(2\pi f_3 k) - \sin(2\pi f_2 k)}{\pi k} \quad (116)$$

$$B = \{b_{1-j}\}$$

$$l_k = 2 \int_0^{f_1} \cos(2\pi f k) df = \frac{\sin(2\pi f_1 k)}{\pi k} \quad (117)$$

$$L = \{l_{i-j}\}$$

The selection matrix  $S$ , (introduced earlier also) is defined as

$$S = [0 \mid I_m \mid 0] \quad (118)$$

where  $I_m$  is an  $m \times m$  identity matrix. The matrix selects the observed  $m$  samples from an infinite size vector. The observation vector can be written as

$$y = Ss + Sc + n. \quad (119)$$

Since  $s$  and  $c$  are bandlimited signals they satisfy  $Bs = s$  and  $Lc = c$ .

Thus, we can write

$$y = SBs + SLc + n \quad (120)$$

$$y = S[B \ L] \begin{bmatrix} s \\ c \end{bmatrix} + n$$

$$\triangleq Hx + n \quad (121)$$

when  $H = S[B \ L]$  and  $x = \begin{bmatrix} s \\ c \end{bmatrix}$ .

Now the problem is to find an estimate of  $x$ , given  $y$ . In the absence of clutter ( $c=0$ ), the problem reduces to the extrapolation problem considered earlier. The solution  $\hat{x}$  obtained by minimizing the least squares norm

$$J = \|y - Hx\|^2$$

is given by

$$\hat{x} = (H^T H)^{-1} H^T y$$

where  $(H^T H)^{-1} H^T$  is the pseudo inverse operator of  $H$  provided  $(H^T H)^{-1}$  exists. However, this is not the case as explained below.

The infinite matrices  $B$  and  $L$  are Toeplitz and are diagonalized by the Fourier transform operator. Therefore,

$$F[B \ L]F^T = \Lambda(f)$$

where  $F$  is the Fourier transform operator and  $F^T$  is its conjugate transpose and  $\Lambda(f)$  is as shown in figure 11.

This also implies that the operators  $B$  and  $L$  are idempotent i.e.,  $B^2 = B$  and  $L^2 = L$ . From Fig. 11 the operator  $[B \ L]$  is singular which in turn implies

$$H^T H = \begin{bmatrix} B^T \\ L^T \end{bmatrix} S^T S [B \ L] \quad (123)$$

is singular, and has a rank of atmost  $m$ .

A number of approaches are once again possible to find the pseudo inverse of  $H$ , as discussed in the previous section. Here we consider a two step gradient method.

An example of a two step gradient method where initial convergence is extremely rapid is the Conjugate Gradient Method [12,13] which is based on the following ideas. Let  $Q$  denote the matrix  $H^T H$  (which has dimension  $2N \times 2N$  to extrapolate the signal and clutter each to  $N$  points).

Let  $\ell = 2N$ . A set of  $\ell$  vectors  $\{d_i\}$  are conjugate if

$$d_i^T Q d_j = 0, \quad i \neq j.$$

Since  $Q$  is symmetric a set of such vectors exist and forms a basis. The solution vector can, therefore, be written as

$$\hat{x} = \sum_{i=1}^{\ell} \alpha_i d_i .$$

The scalars  $\{\alpha_i\}$  and vectors  $\{d_i\}$  must be found in a computationally feasible way. One way is via the algorithm [13]

$$\begin{aligned} \hat{x}^{k+1} &= \hat{x}^k + \alpha_k d_k \\ \alpha_k &= \frac{-C_k^T d_k}{d_k^T Q d_k} \end{aligned} \quad (124)$$

where the  $\{d_k\}$  are generated by

$$\begin{aligned} d_{k+1} &= -C_{k+1} + \beta_k d_k \\ \beta_k &= \frac{C_{k+1}^T Q d_k}{d_k^T Q d_k} . \end{aligned} \quad (125)$$

The vectors  $\{C_k\}$  are the gradients of  $J$  at each iteration

$$C_k = Q \hat{x}^k - H^T y = C_{k-1} + \alpha_{k-1} Q d_{k-1} . \quad (126)$$

and the initial conditions are

$$\hat{x}^1 = y \quad \text{and} \quad d_1 = C_1 . \quad (127)$$

The minimum is achieved in at most  $\ell$  steps, and the method is step for step better than a gradient method. A very attractive feature of this

algorithm is that large reductions in error are achieved in the first few steps [13].

Looking at the computations involved, we see that except for the single matrix vector product  $Qd_k$ , all vector operations involve only  $\ell$  multiplications. Hence, the order of computations for each iteration is  $4mN$ .

(Note that  $Q$  is composed from Toeplitz matrices, so that advantage of FFT method could be taken to evaluate  $Qd_k$ .) When  $Q$  is highly ill-conditioned, or singular, the iterations must be stopped at an optimum point. Alternatively, we may add a small value  $\epsilon$  of the order of  $10^{-6}$  to each diagonal term of  $Q$ . This minimizes convergence ambiguity due to ill-conditioning and stabilizes the iterations greatly.

# VI. A MEAN SQUARE EXTRAPOLATING FILTER

With the formulation and notation of the problem of Section V, we have<sup>†</sup>

$$y = Hx + n. \quad (128)$$

Now we assume that  $x$  is a random Gaussian vector whose autocorrelation matrix is denoted by  $R_x$ . The minimum mean square estimate of  $x$  is given by the Wiener filter as\*

$$\hat{x} = [E(xy^T)][E(yy^T)]^{-1}y \triangleq Gy. \quad (129)$$

Assuming noise to be independent of  $x$ , it is easy to obtain

$$\begin{aligned} G &= R_x H^T (H R_x H^T + R_n)^{-1} \\ R_n &= E(n n^T) \end{aligned} \quad (130)$$

which is equivalent to the equations

$$\begin{aligned} (H R_x H^T + R_n)z &= y \\ \hat{x} &= R_x H^T z \\ R_x &= E[x x^T] = E\left[\begin{pmatrix} s \\ c \end{pmatrix} \begin{pmatrix} s^T & c^T \end{pmatrix}\right] \\ &= \begin{bmatrix} E(ss^T) & E(sc^T) \\ E(cs^T) & E(cc^T) \end{bmatrix}. \end{aligned} \quad (140)$$

<sup>†</sup> We note here the similarity between the extrapolation problem and the restoration problem in image processing [17]. For example, an image blurred by a low pass type operator  $H$  and contaminated by additive noise would give rise to an equation similar to (128).

\* Here  $E$  is the expectation operator.

If signal and clutter are uncorrelated (non overlapping power spectra)

$$R_x = \begin{bmatrix} E(ss^T) & 0 \\ 0 & E(cc^T) \end{bmatrix}. \quad (141)$$

In the special case when the noise and clutter are absent, we have

$$H = B$$

$$R_n = 0.$$

Thus, we have  $z = (SBR_x B^T S^T)^{-1} y$ . Since  $s(k)$  is a bandlimited random process, we must have

$$BR_x B^T = R_x$$

which gives

$$\hat{x} = R_x S^T (S R_x S^T)^{-1} y.$$

In the worst case, when we do not know  $R_x$ , we can simply assume the power spectrum of  $x(k)$  to be flat in its bandwidth so that

$$R_x = B$$

and  $\hat{x} = BS^T(SBS^T)^+ y$  where  $(SBS^T)^+$  is the pseudo inverse of  $(SBS^T)$ . This is the same result as obtained by Cadzow [See Eqn. (95)]. Thus Cadzow's one shot method is a special case of this extrapolation filter. In the presence of noise, the extrapolation filter estimate is

$$\hat{x} = BS^T(SBS^T + R_n)^{-1} y$$

where  $(SBS^T + R_n)^{-1}$  exists and is unique.

## VII. A RECURSIVE EXTRAPOLATION ALGORITHM

In this section we present a recursive least squares algorithm based on Kalman filtering techniques where the extrapolated signal estimate is updated recursively as a new observation sample arrives.

Based on the formulation of the problem as in section V we rewrite equation (121):

$$y = Hx + n$$

The  $k^{\text{th}}$  observation  $y_k$  can be written as

$$y_k = h_k^T x + n_k \quad \text{and} \quad k=0,2,\dots,m-1 \quad (142)$$

where  $h_k^T$  is the  $k^{\text{th}}$  row of  $H$  and  $n_k$  is zero mean white Gaussian noise.

The state equation for the unknown extrapolated vector  $x$  can be written as

$$x_{k+1} = x_k \quad (143)$$

with initial condition  $x_0 = x$ , where  $x$  is a random vector whose covariance is given by

$$P_0 = \text{cov}(x_0) = \text{cov}(x) = H \triangleq \{\alpha_{k-l}\} \quad (144)$$

Since  $H$  is idempotent i.e.  $HH = H$ , it is easy to verify that

$$h_k^T h_{k-l} = \alpha_{k-l} \quad (145a)$$

$$H h_k = h_k \quad (145b)$$

and

$$h_k^T H = h_k^T \quad (145c)$$

The Kalman filter associated with equations (142)-(144) is the recursive least squares filter (e.g. see Nahi [16])

$$\hat{x}_{k+1} = \hat{x}_k + g_k (y_k - h_k^T \hat{x}_k), \quad \hat{x}_0 = 0 \quad (146)$$

where  $\hat{x}_k$  is the  $k^{\text{th}}$  estimate of  $x$  and  $g_k$  is the Kalman filter gain.

The associated Riccati equation is

$$P_{k+1} = (I - \frac{1}{c_k} P_k h_k h_k^T) P_k (I - \frac{1}{c_k} h_k h_k^T P_k) + g_k \sigma_n^2(k) g_k^T \quad (147)$$

and

$$c_k = h_k^T P_k h_k + \sigma_n^2(k), \quad \sigma_n^2(k) = E[n_k^2] \quad (148)$$

$$g_k = \frac{1}{c_k} P_k h_k \quad (149)$$

Equation (147) then reduces to

$$P_{k+1} = P_k - \frac{1}{c_k} P_k h_k h_k^T P_k \quad (150)$$

Letting  $P_k = H + \partial P_k$ , and using equations (145),

$$\partial P_{k+1} = \partial P_k - \frac{1}{c_k} (I + \partial P_k) h_k h_k^T (I + \partial P_k) \quad (151)$$

Defining

$$\xi_{k,l} = \partial P_k h_l \quad (152a)$$

and

$$\beta_{k,l} = \xi_{k,k}^T h_l \quad (152b)$$

it follows from (151) that

$$\xi_{k+1,l} = \xi_{k,l} - \frac{1}{c_k} [\alpha_{l-k} + \beta_{k,l}] (h_k + \xi_{k,k}); k=0, \dots, m-1; l=k, k+1, \dots, m-1; \quad (153)$$

$$\xi_{0,l} = 0, \forall l$$

From (149) it follows that

$$g_k = \frac{1}{c_k} (h_k + \xi_{k,k}) \quad (154)$$

From (145), (148) and (152), we get

$$c_k = \alpha_0 + \sigma_n^2(k) + \beta_{k,k} \quad (155)$$

Assuming that we are basing our extrapolation on  $m$  observed sample values and that we are extrapolating to  $N$  points, the maximum storage required is  $mN$ .

The major computation is in the calculation of the  $\xi_{k,l}$  (153) and it is easily seen that the order of computations involved is  $\sim O(m^2N)$ .

# VIII. EXAMPLES, RESULTS AND COMPARISONS

We will now consider several examples to study the performance of the foregoing algorithms. In Example 1 below, we consider the observations to be of the form

$$y(k) = s(k) + c(k) + n(k) , \quad -M \leq k \leq M$$

where  $c(k)$  is the bandlimited clutter sequence, and  $s(k)$  is a bandlimited signal whose band limits are known, i.e.,

$$S(\omega) = 0, \omega \notin [\omega_1, \omega_2].$$

$n(k)$  is a zero mean white noise process with variance  $\sigma_n^2$ .

In examples 2-9 the observations are of the form

$$y(k) = s(k) + n(k) ; \quad -M \leq k \leq M .$$

The following examples have been considered.

1a.  $y(k) = s(k) + c(k) + n(k) ; \quad -8 \leq k \leq 8$

$$s(k) = 1.69 \sin(.39\pi k)$$

$$\omega_1 = 0.382\pi$$

$$\omega_2 = 0.397\pi$$

$$\text{clutterband: } \omega_c = [0, 0.16\pi]$$

$$\sigma_n^2 = 0.13$$

$$\text{SCR} = -4.1\text{dB}, \text{ SNR} = 19\text{dB}$$

1b.  $y(k) = s(k) + c(k); \quad -8 \leq k \leq 8$

$$\omega_1 = .3\pi$$

$$\omega_2 = .4\pi$$

$$\omega_s = .875\omega_2$$

$$s(k) = 1.69\sin(\omega_s k)$$

$$\sigma_n^2 = 0$$

$$2. \quad s(k) = \sin(.99\omega_2 k) + \sin(.85\omega_2 k)$$

$$\sigma_n^2 = 0$$

$$\omega_1 = 0.8\omega_2$$

$$\omega_2 = 2\pi/50$$

$$3. \quad s(k) = \sin(.99\omega_2 k) + \sin(.85\omega_2 k)$$

$$\sigma_n^2 = 0$$

$$\omega_1 = 0$$

$$\omega_2 = 0.1\pi$$

The above two cases have the same signal, but the knowledge of bandlimits is different. In the following cases we have other signals with different bandlimits. In examples 8 and 9 we have additive noise also.

$$4. \quad s(k) = \sin(.365\pi k) + \sin(.385\pi k)$$

$$\sigma_n^2 = 0$$

$$\omega_1 = .3\pi$$

$$\omega_2 = .4\pi$$

$$5. \quad s(k) = \sin(.365\pi k) + \sin(.385\pi k)$$

$$\omega_1 = .35\pi$$

$$\omega_2 = .4\pi$$

$$\sigma_n^2 = 0$$

$$6. \quad s(k) = \sin(.985\omega_2 k) + \sin(.975\omega_2 k)$$

$$\omega_1 = 0.96\omega_2$$

$$\omega_2 = 0.4\pi$$

$$\sigma_n^2 = 0$$

$$7. \quad s(k) = 0.1 \frac{\sin \omega_2 k}{\omega_2 k}$$

$$\omega_1 = 0$$

$$\omega_2 = .1\pi$$

$$\sigma_n^2 = 0$$

$$8. \quad s(k) = 0.1 \frac{\sin \omega_2 k}{\omega_2 k}$$

$$\omega_1 = 0$$

$$\omega_2 = 0.1\pi$$

$$E(n(k)) = 0, \quad \sigma_n^2 = 0.01$$

$$\text{SNR} = 7.4\text{dB}$$

$$9) \quad s(k) = \sin(.99\omega_2 k) + \sin(.85\omega_2 k)$$

$$\omega_1 = .8\omega_2$$

$$\omega_2 = 2\pi/50$$

$$E(n(k)) = 0$$

$$\sigma_n^2 = 0.1, \quad \text{SNR} = 21.6\text{dB}$$

In example 1a, application of an eighth order autoregressive model to the given data yields the spectrum shown in Fig 1(e). The estimate seems

to be completely dominated by the clutter. A 256-point DFT-spectrum estimate of the data is shown in Fig. 1(f), and clearly, this also is most unsatisfactory. Figures 1(g) and 1(h) show the signal extrapolated to 199 points by the Conjugate Gradient Algorithm (using only 10 iterations) and the Mean Square Extrapolation Filter. Figures 1(i) and 1(j) show the corresponding results for the extrapolated clutter. The Maximum Entropy method is applied to the extrapolated signal and extrapolated clutter separately using a fifteenth order model. Fig. 1(k) and 1(l) show the extrapolated signal spectra using the Conjugate Gradient Algorithm and the M.S. Extrapolation filter. They yield the signal peak at the correct location. Figures 1(m) and 1(n) show the corresponding clutter spectra.

In example 1b, the signal bandwidth is increased from the previous case. The actual spectra of signal and clutter are shown in fig. 2(a). The FFT spectrum using 256 points is shown in fig. 2(b), and the Maximum Entropy Spectrum using an eight order model in fig. 2(c). After extrapolating signal and clutter to 125 points each, using the conjugate gradient algorithm their respective spectra are calculated by the Maximum Entropy Method using a 15<sup>th</sup> order model. This is shown in figures 2(d) and 2(e) respectively. Figure 2(d) shows a peak at the signal frequency along with two subsidiary peaks which are seen to occur exactly at the filter cut-off frequencies.

In examples 2-9 we consider either signal only or signal with additive, white, Gaussian noise.

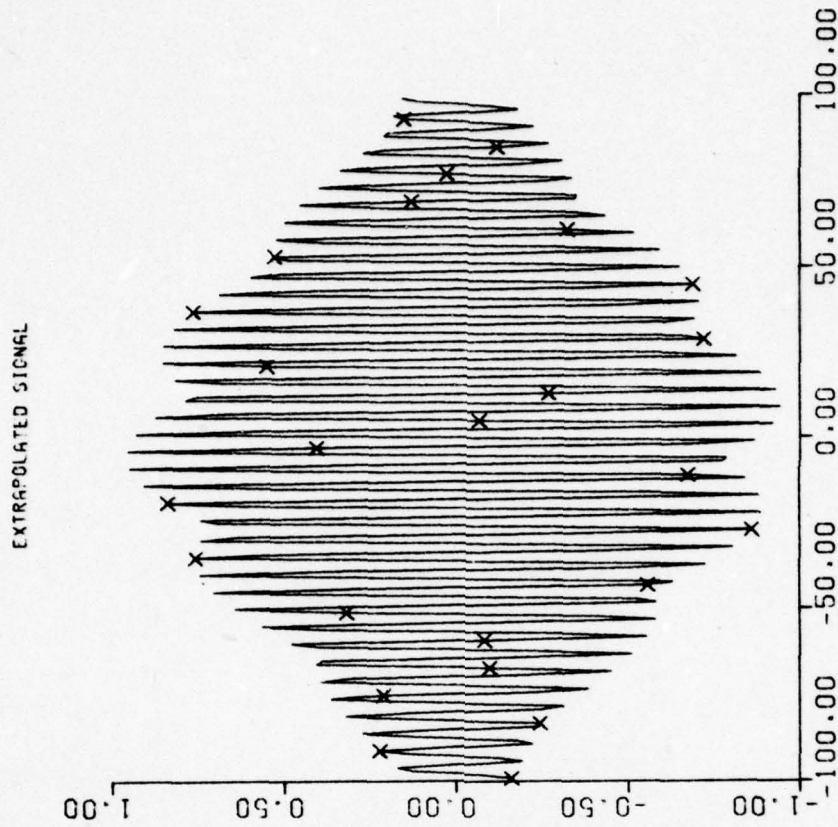


Fig. 1h: Signal Extrapolated by M.S. Extrapolation Filter

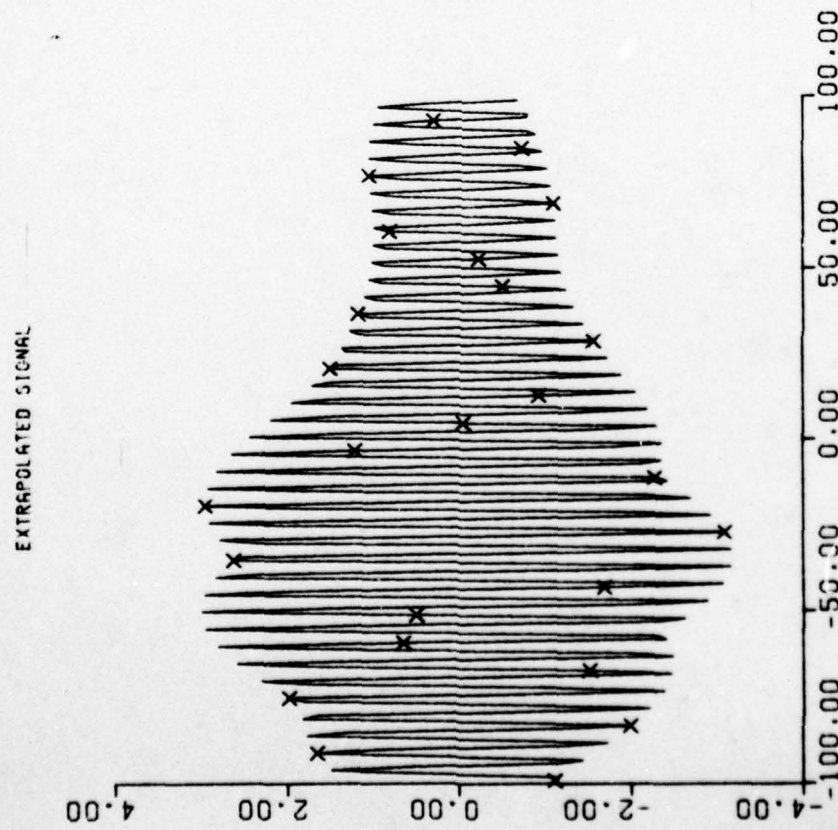


Fig. 1g: Signal Extrapolated by Conjugate Gradient Algorithm

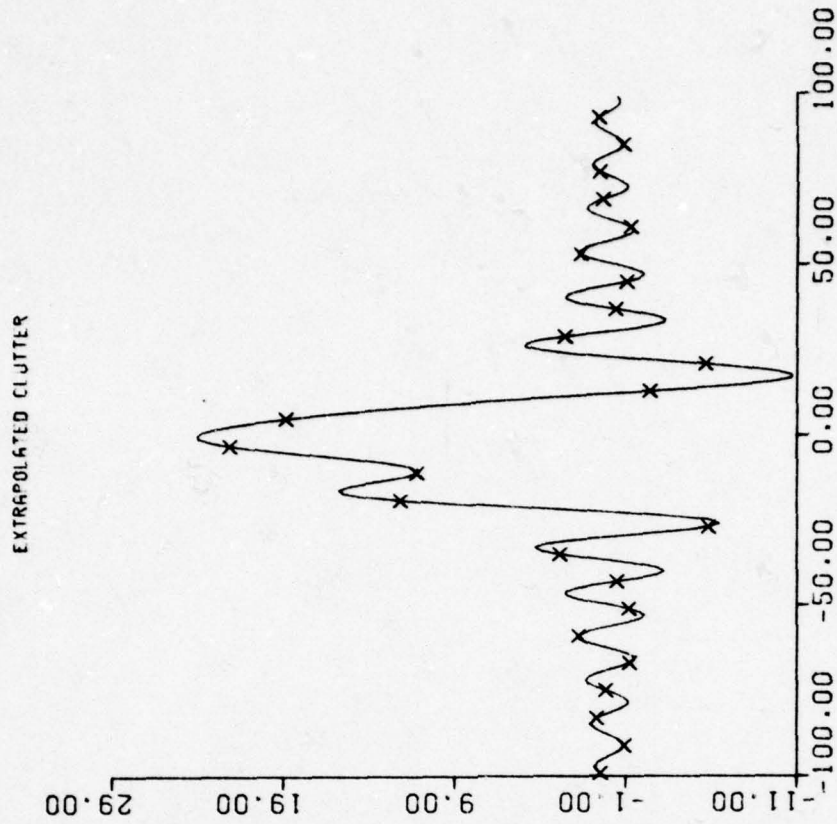


Fig. 1j: Clutter Extrapolated by M.S. Extrapolation Filter

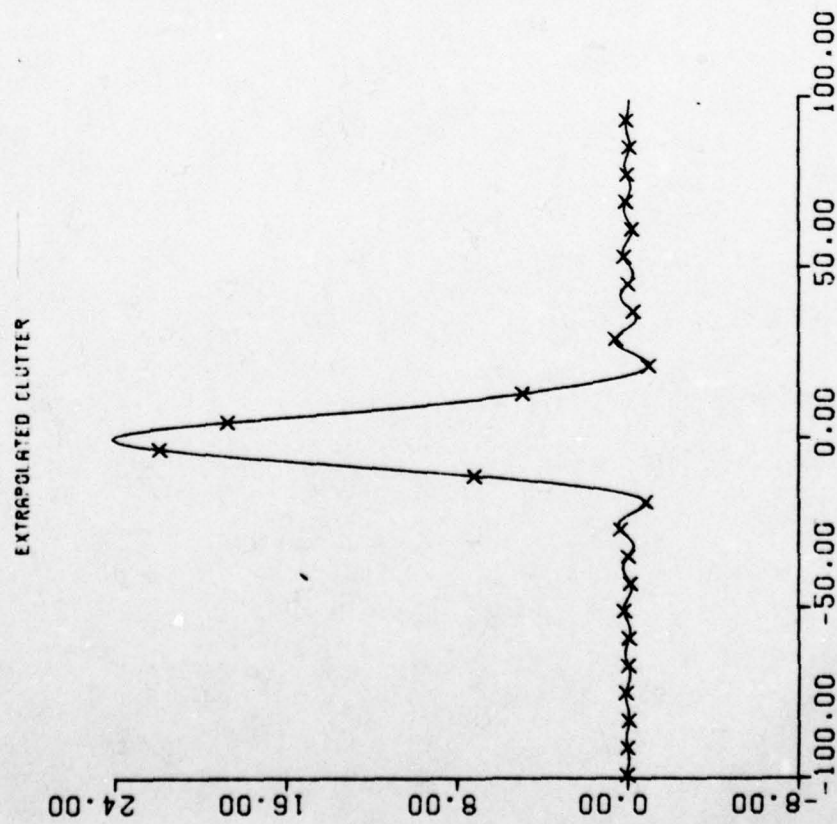


Fig. 1i: Clutter Extrapolated by Conjugate Gradient Algorithm

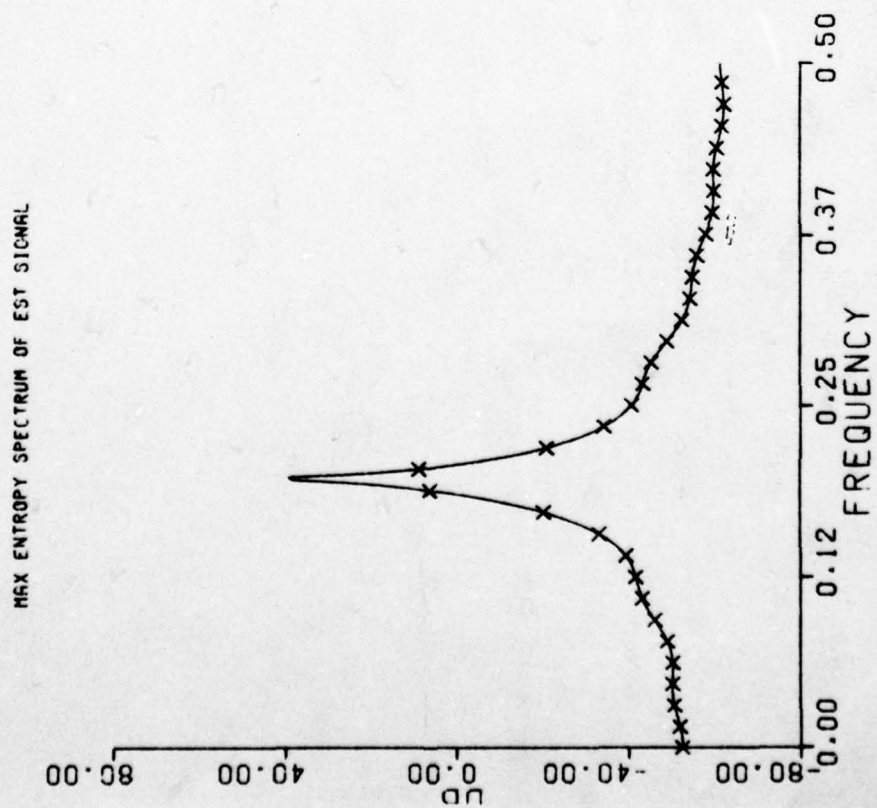
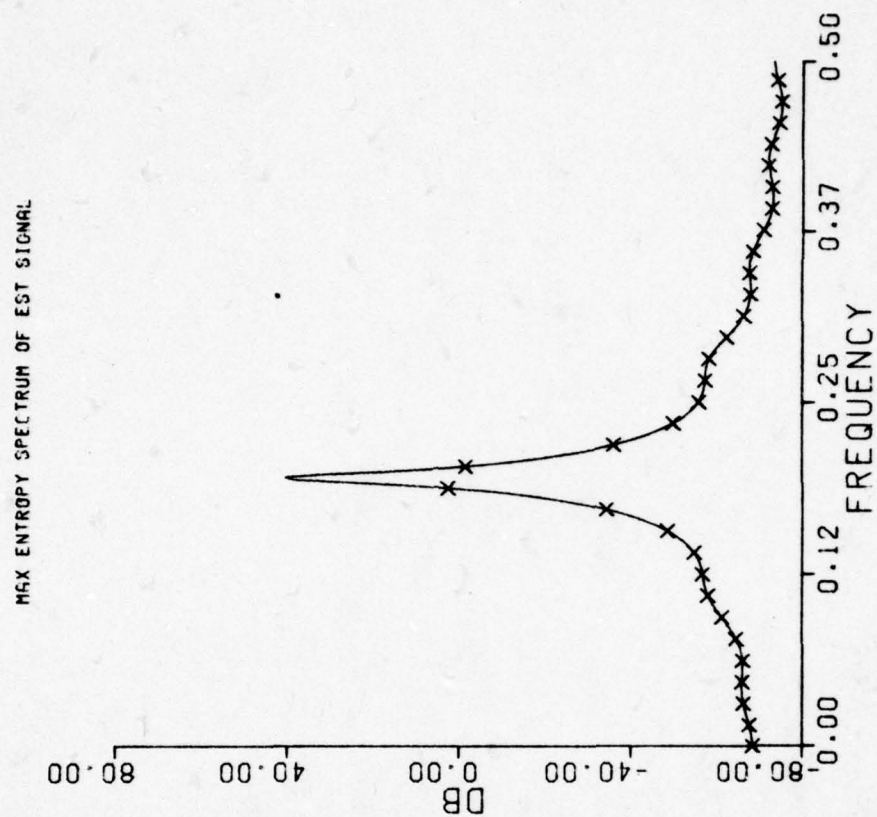


Fig. 11: Max Entropy Spectrum of (1h) (15th order model)

Fig. 1k: Max Entropy Spectrum of (1g) (15th order model)

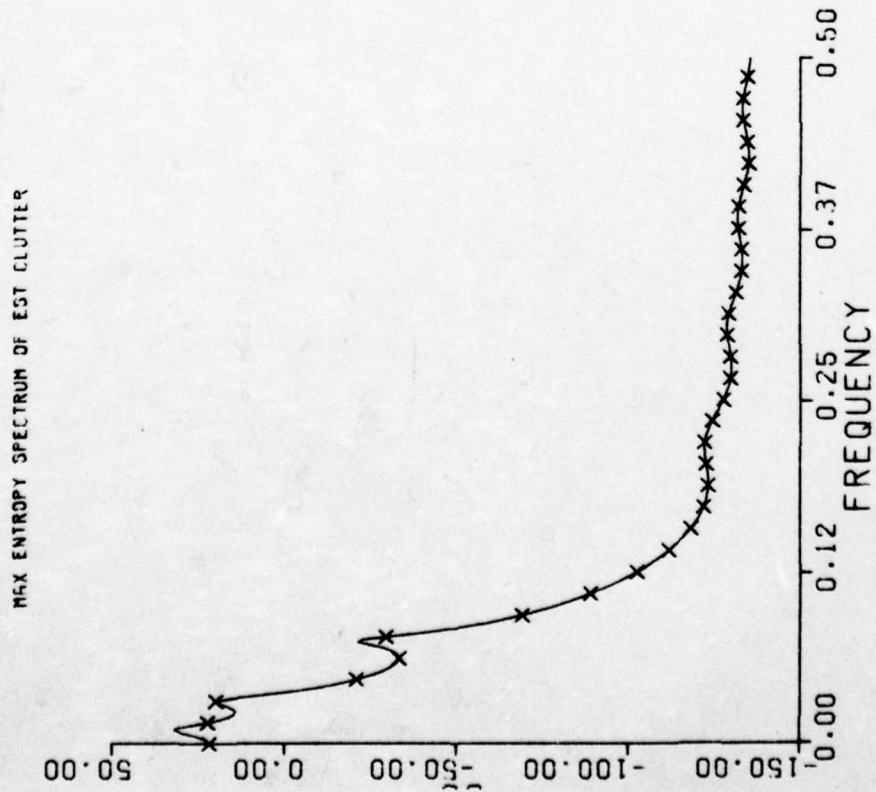
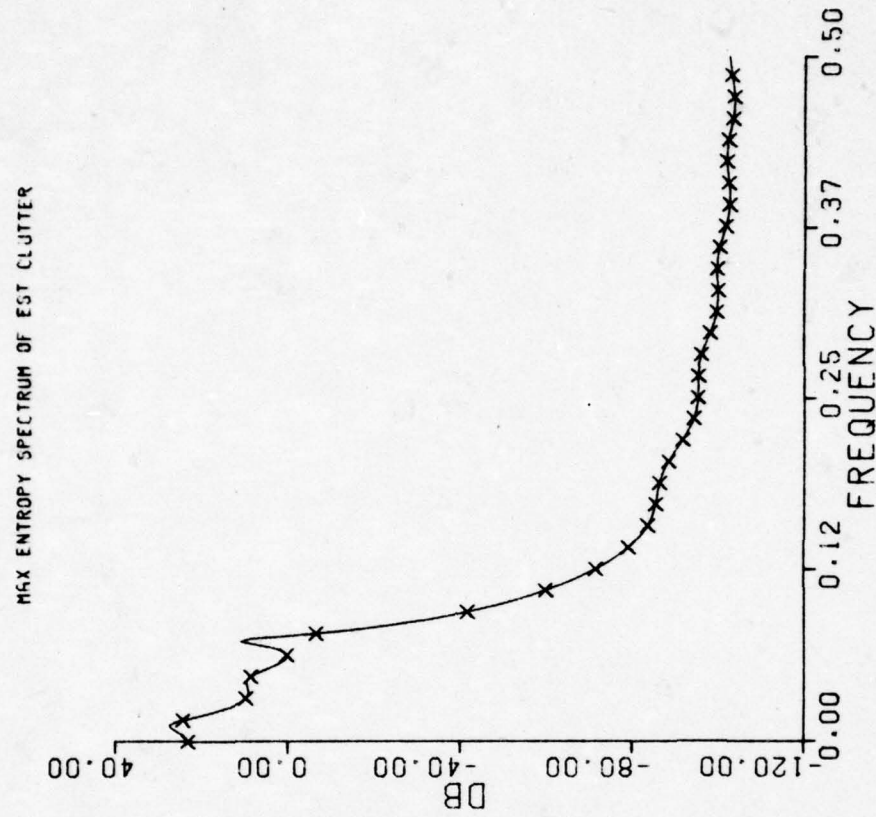


Fig. 1m: Max Entropy Spectrum of (1i) (15th order Model)      Fig. 1n: Max Entropy Spectrum of (1j) (15th order model)

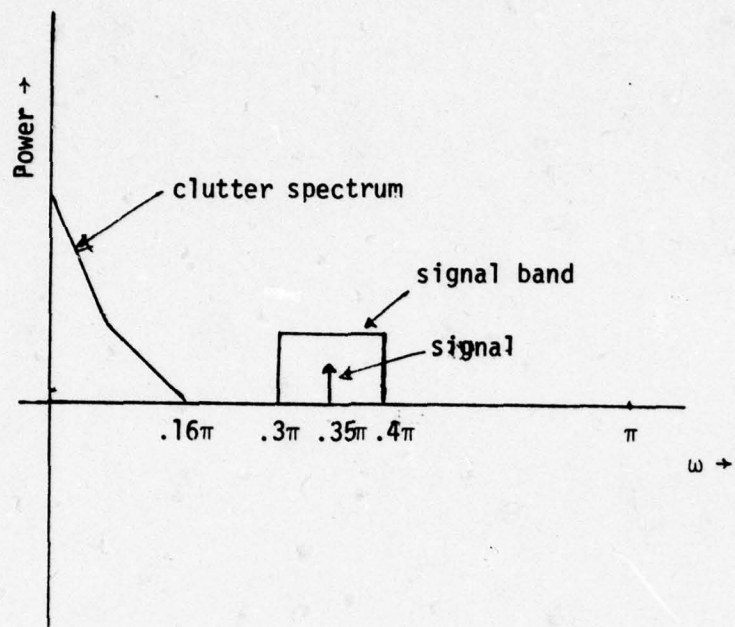


Fig. 2a: Spectra of Signal and Clutter

Fig. 2b: 256 pt. Power Spectrum of Observations by FFT

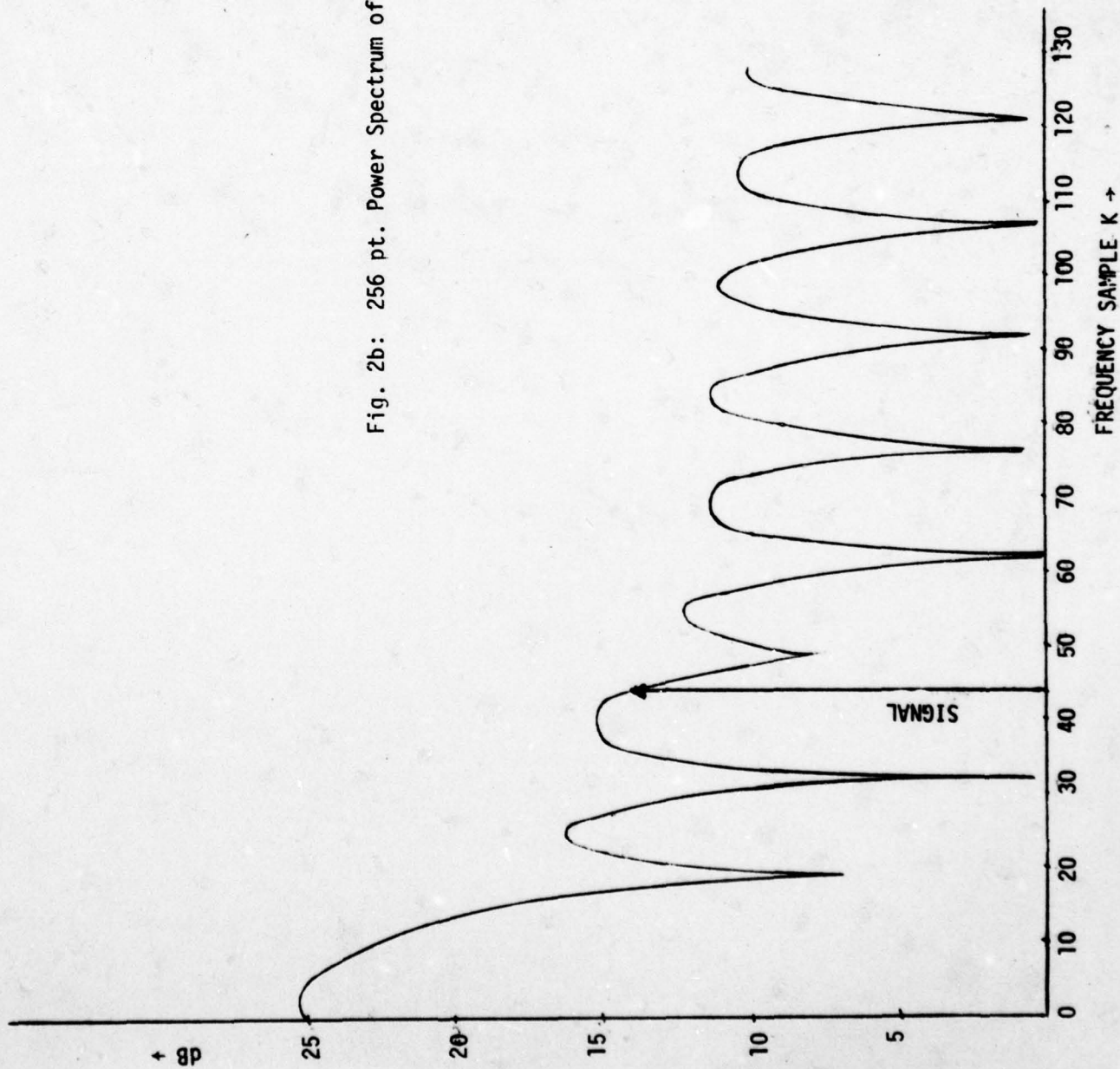


Fig. 2c: Max Entropy Spectrum of Observations by an 8th Order Model

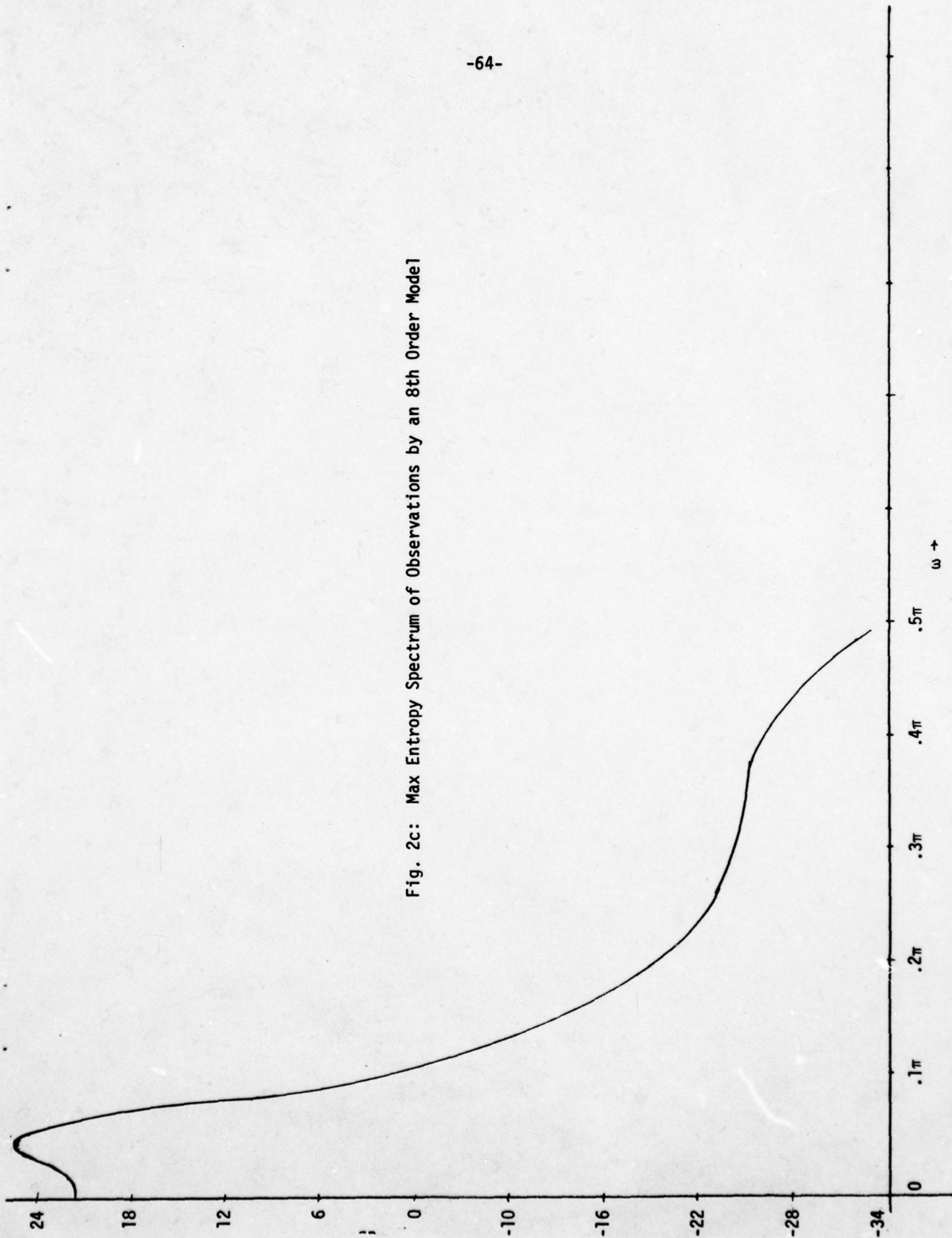


Fig. 2d: Max Entropy Spectrum of Signal  
Extrapolated to 125 pts. (Ex. 2)

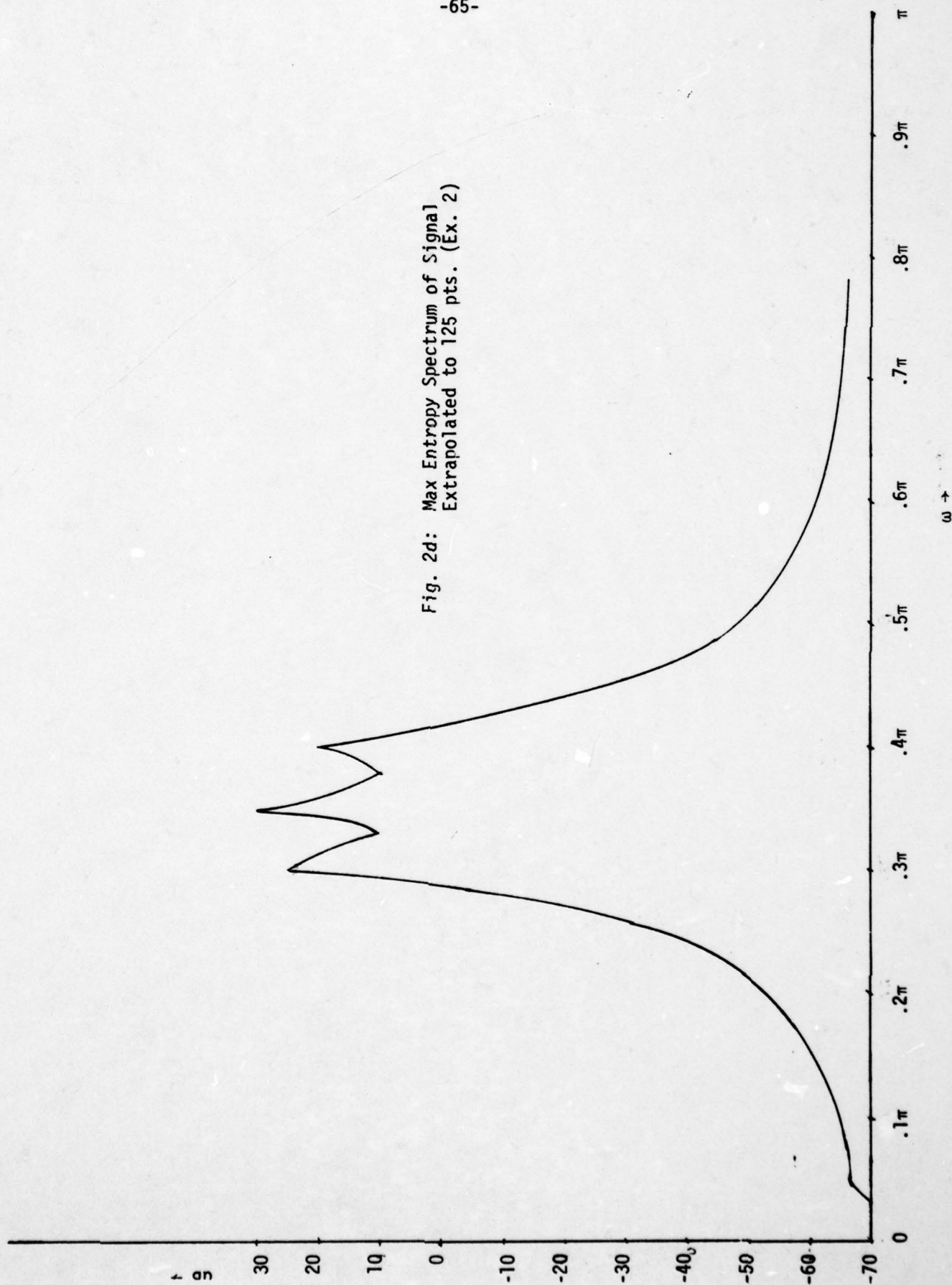
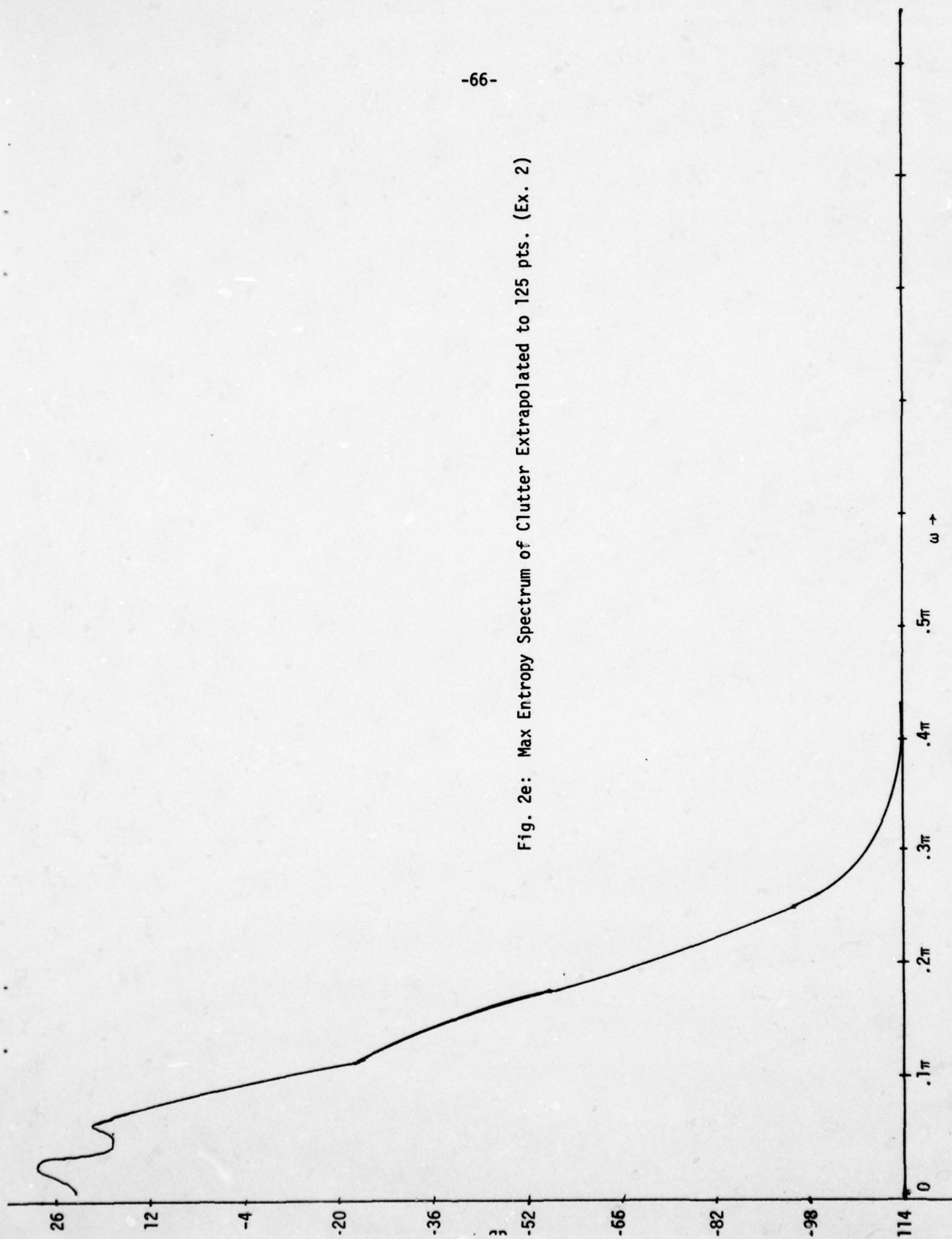


Fig. 2e: Max Entropy Spectrum of Clutter Extrapolated to 125 pts. (Ex. 2)



The observed data consists of 17 samples and it is extrapolated to 199 points using

1. Papoulis' iterative algorithm
2. the matrix  $E_C$
3. the matrix  $E_C$  after it has been stabilized
4. the conjugate gradient algorithm
5. the M.S. extrapolation filter for Examples 8 and 9 when the signal contains noise.

Extrapolation using Papoulis' algorithm was done with 30 iterations and with 10 iterations using the conjugate gradient algorithm.

Since all the algorithms ultimately yield a minimum norm least-squares solution, we expect equivalent results from all the algorithms. Figure 3(c) shows the extrapolation using Papoulis' algorithm, and is seen to give a reasonable result. Figure 3(d) shows the extrapolation via the matrix  $E_C$ . After a stabilizing diagonal term of the order of  $10^{-6}$  is added to the matrix  $(SLS^T)$ , the extrapolation obtained is as shown in figure 3(e). A close examination of figures 3(a), 3(d) and 3(e) shows a slight phase shift of the signal 3(d). This may be attributable to the ill-conditioning of  $(SLS^T)$ . Figure 3(f) shows the extrapolation by the Conjugate Gradient method after 10 iterations. This algorithm gives a slightly inferior extrapolation compared to the others in this case, and perhaps needs a few more iterations. In later examples, however, it is seen that 10 iterations are sufficient.

In the set of figures 4, we have the same signal and observations, but the uncertainty in bandwidth is increased over the previous case. The relatively degraded extrapolations achieved now show that there is a trade-off between bandwidth uncertainty and extrapolation length, which is to be expected. The set of figures 5, 6, 7 lead to similar conclusions for

signals with differing frequencies and different bandwidths.

In the set of figures 8, we consider a  $\sin x/x$  type of signal. The phase distortions of the extrapolated estimate in figure 8(d) dramatically illustrate the ill-conditioned nature of  $(SLS^T)$ . In the above examples there was no noise in the observations. Now, we consider (examples 9 and 10) signals in additive zero mean, white Gaussian noise at the SNR 7.4dB, 21.6dB respectively. The results show that the Mean Square Extrapolation filter, which takes noise statistics into account produces the best extrapolation (Figs. 9(g) and 10(g)) among the algorithms considered experimentally. Similar results are to be expected from the recursive extrapolation algorithm of section VII. Further simulations are required to study the numerical behavior of this algorithm.

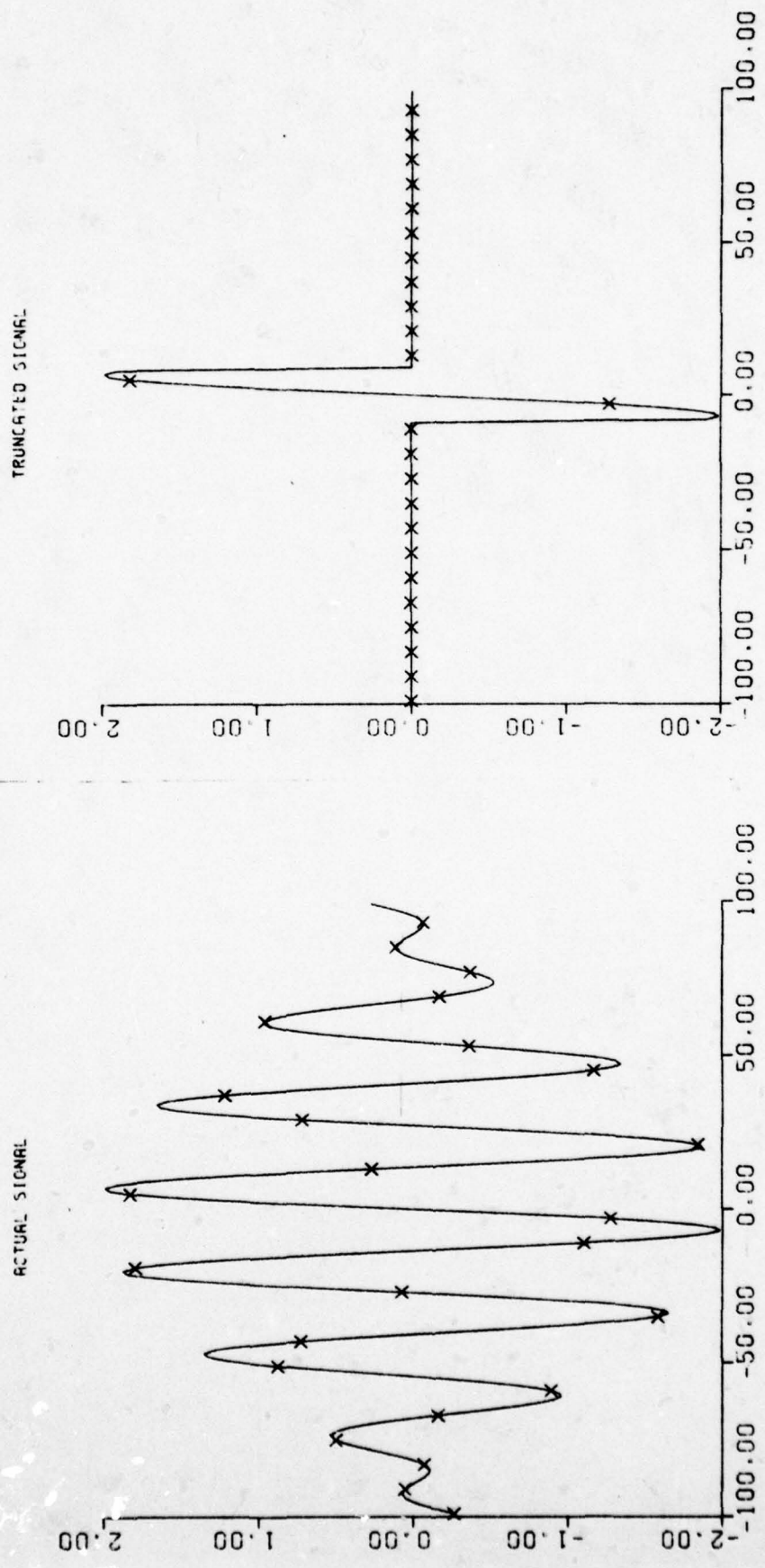


Fig. 3b: Given Observations (17 Samples)

Fig. 3a: Original Signal

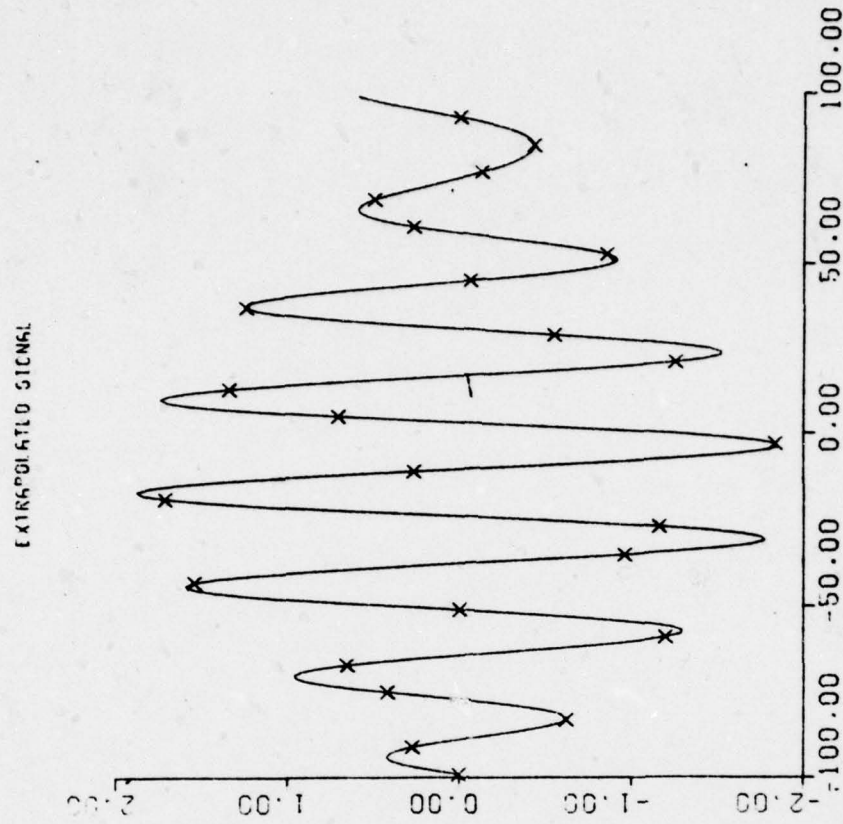


Fig. 3d: Signal Extrapolated Via Matrix  $E_c$

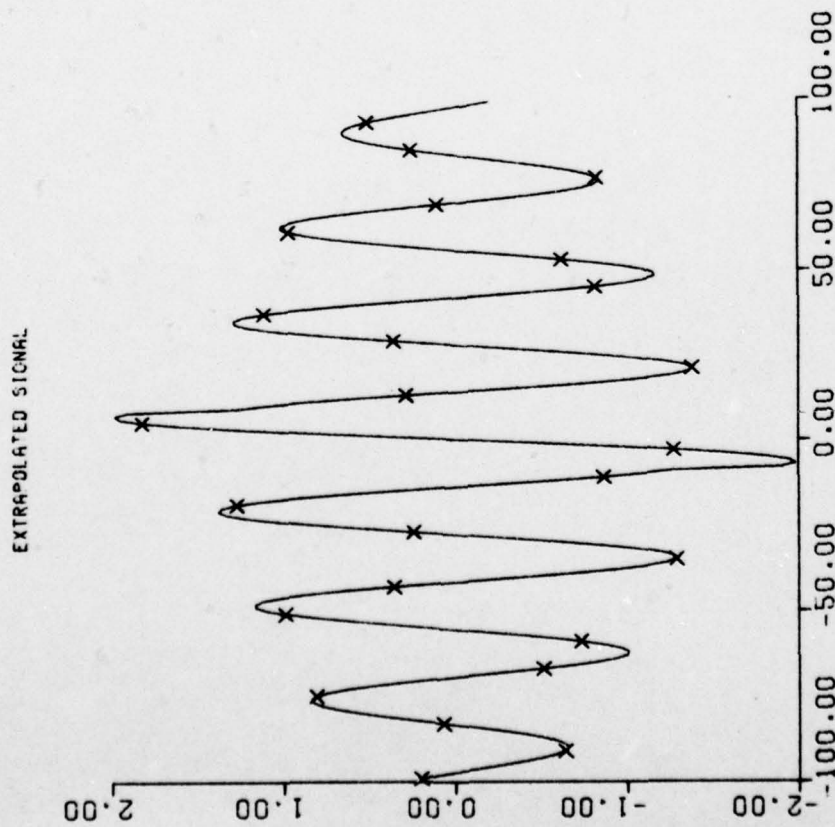


Fig. 3c: Signal Extrapolated by Papoulis' Iterative Algorithm

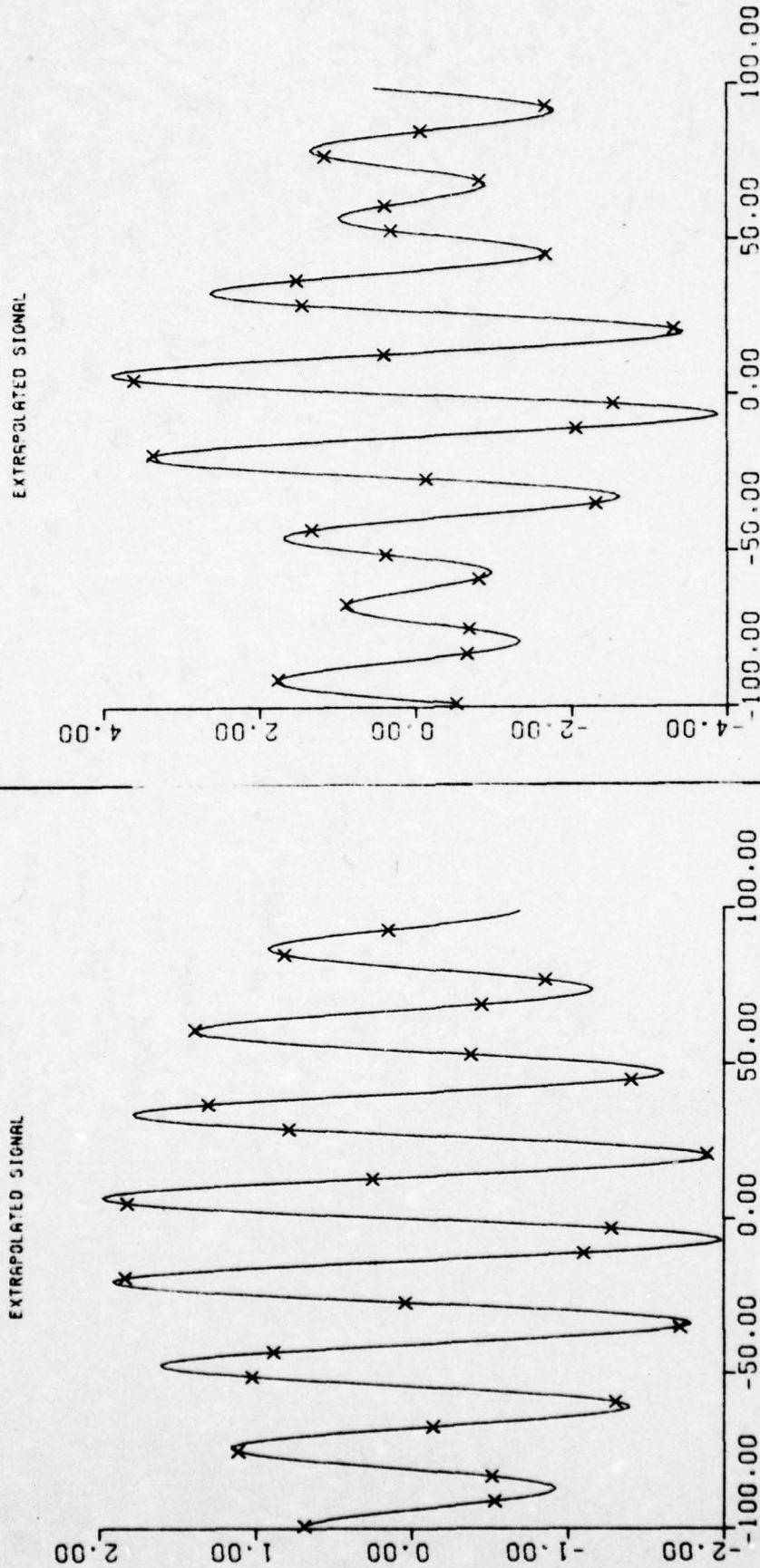


Fig. 3f: Signal Extrapolated by Conjugate Gradient Algorithm

Fig. 3e: Signal Extrapolated After Adding a Stabilizing Diagonal Term to Matrix  $E_c$

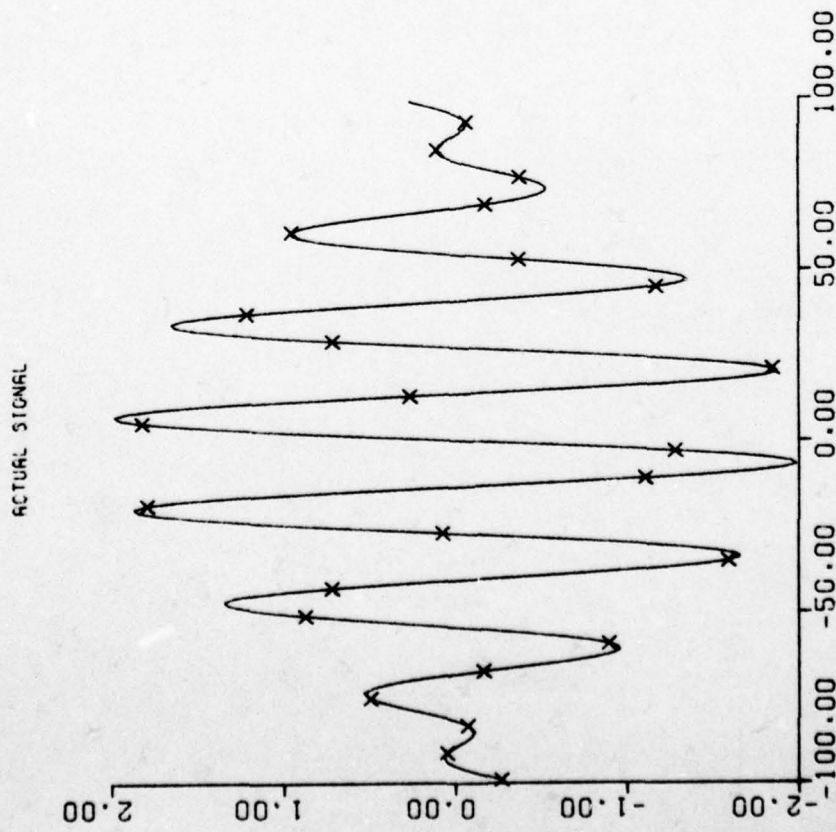
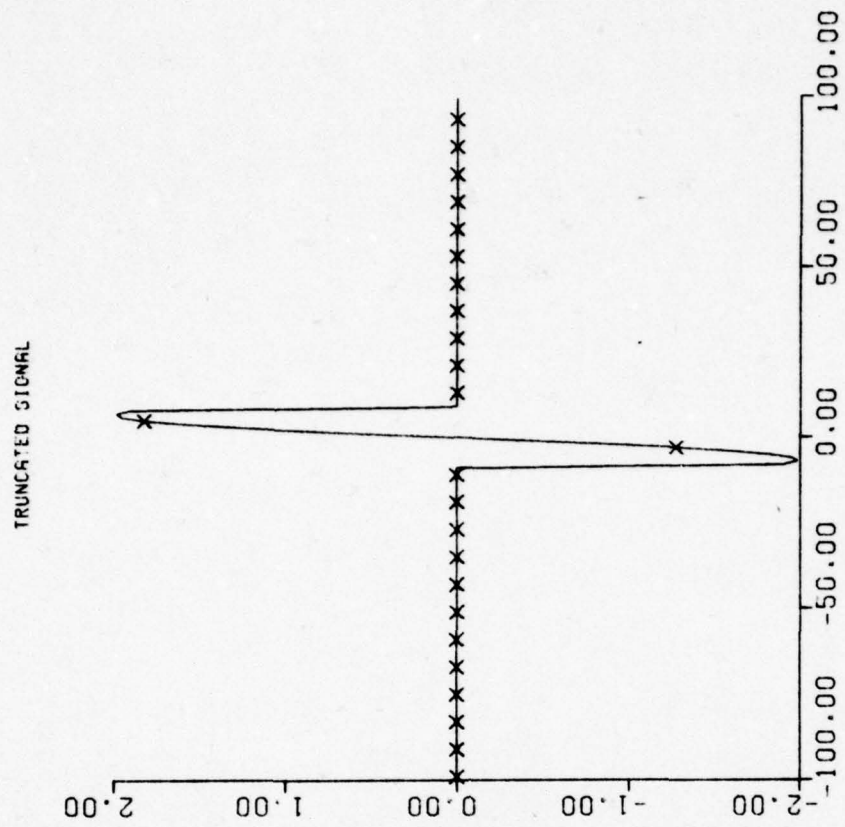


Fig. 4b: Given Observations (17 Samples)

Fig. 4a: Original Signal

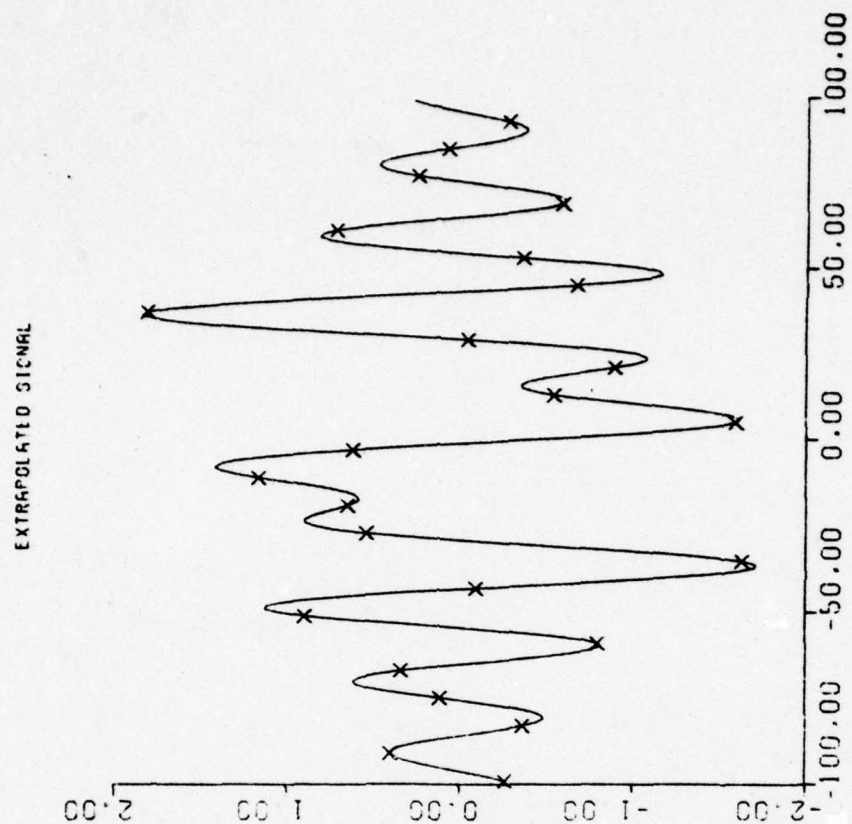


Fig. 4d: Signal Extrapolated Via Matrix  $E_c$

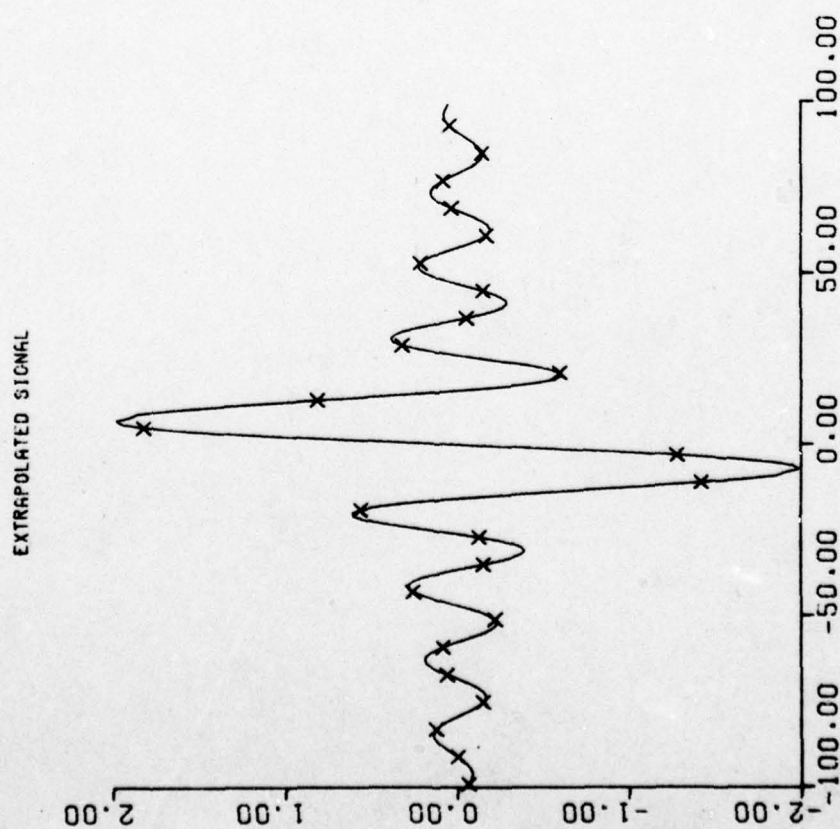


Fig. 4c: Signal Extrapolated by Papoulis' Iterative Algorithm

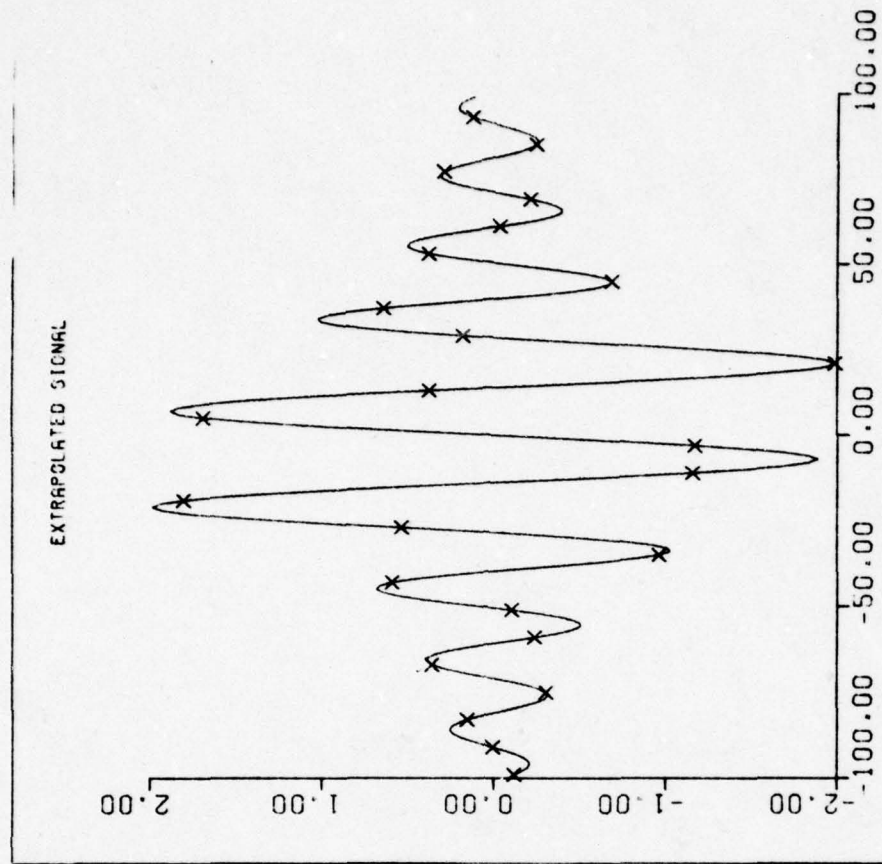


Fig. 4e: Signal Extrapolated After Adding a Stabilizing Diagonal Term to Matrix  $E_c$

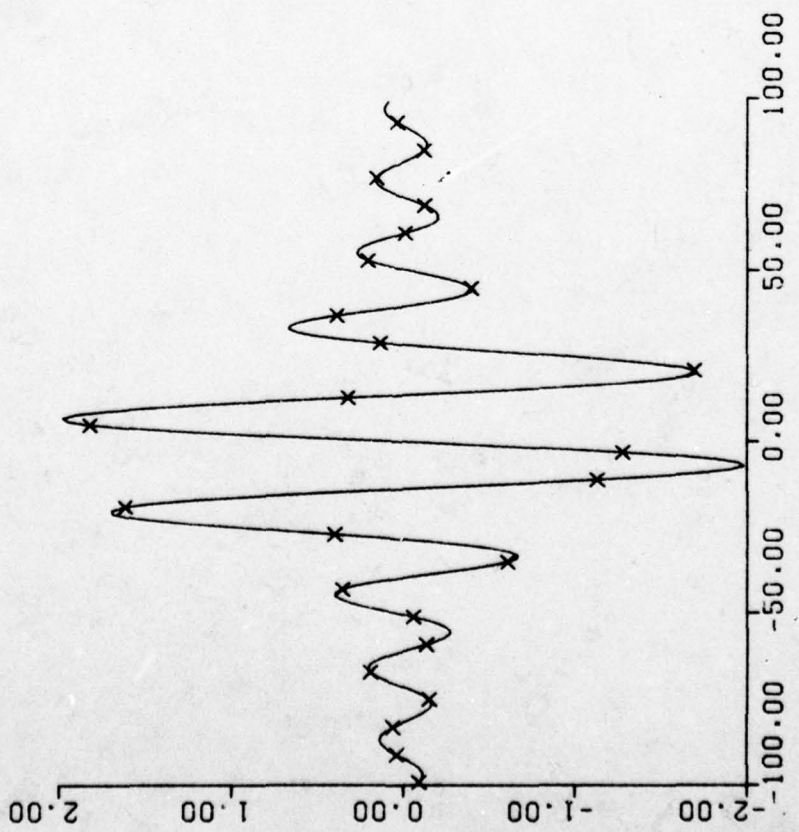


Fig. 4f: Signal Extrapolated by Conjugate Gradient Algorithm

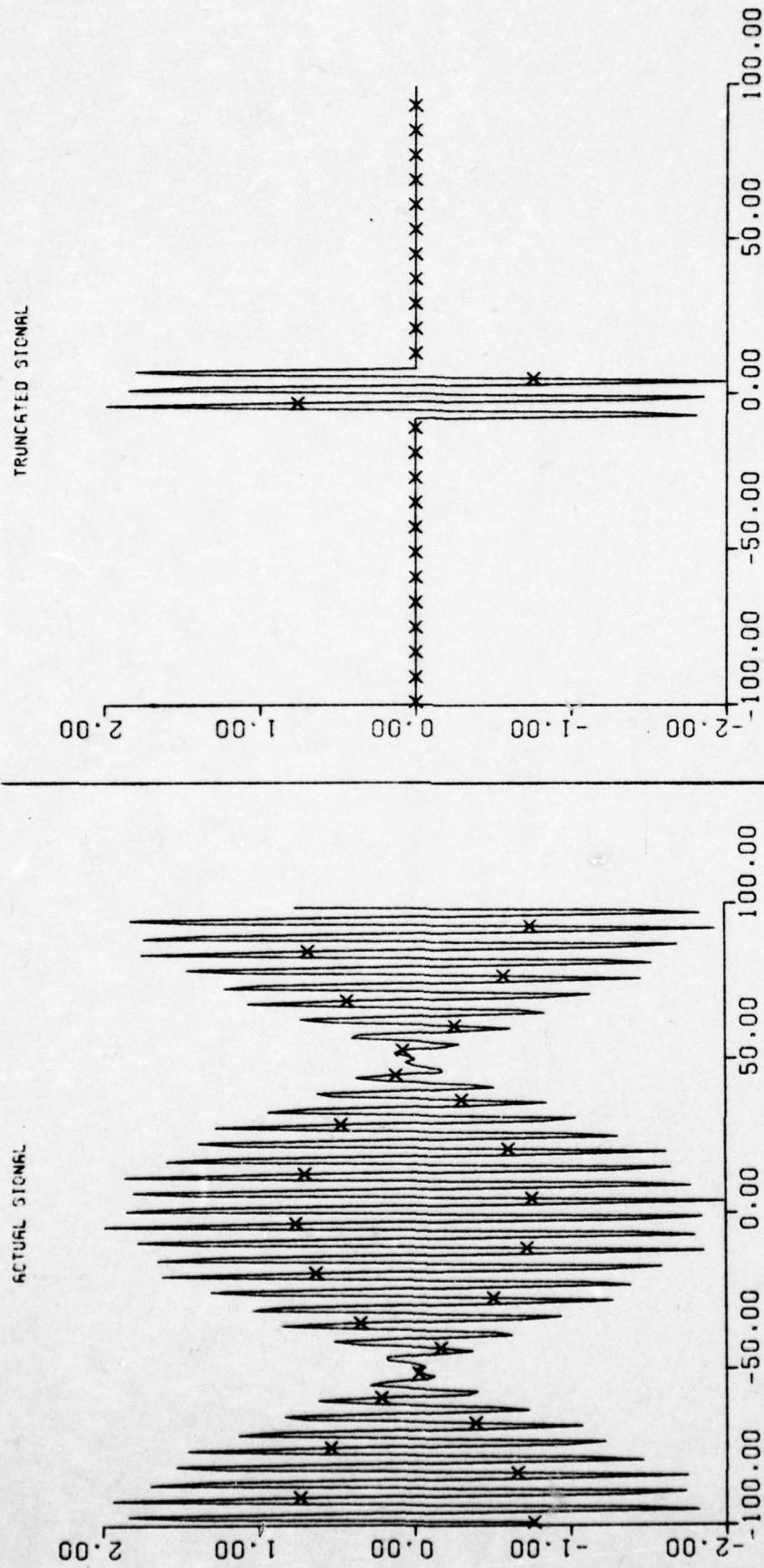


Fig. 5b: Given Observations (17 Samples)

Fig. 5a: Original Signal

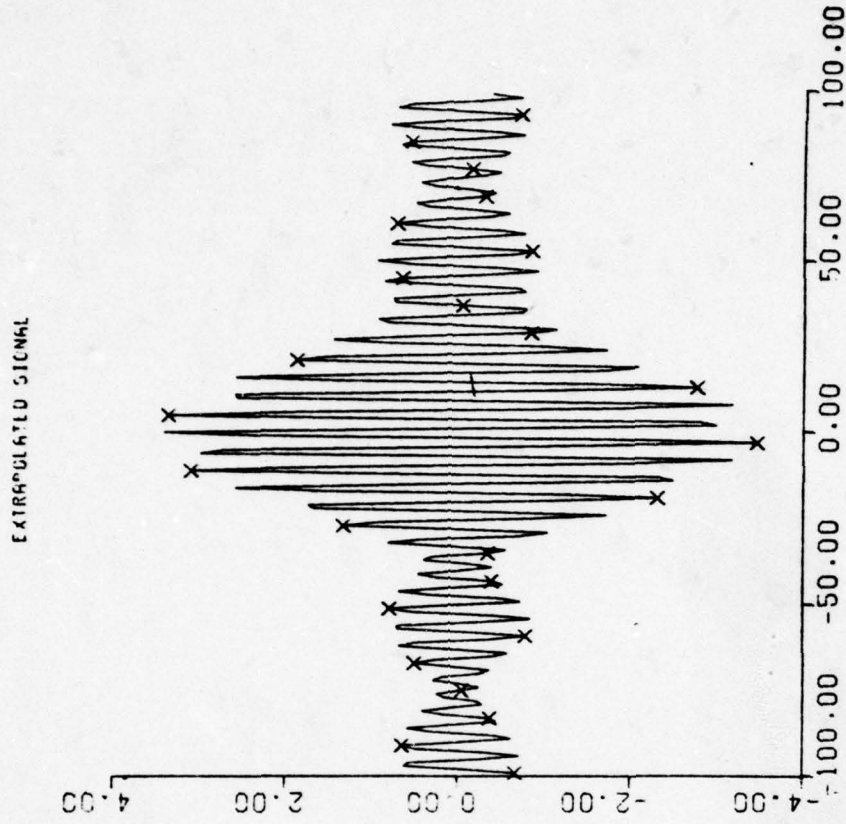


Fig. 5d: Signal Extrapolated Via Matrix  $E_c$

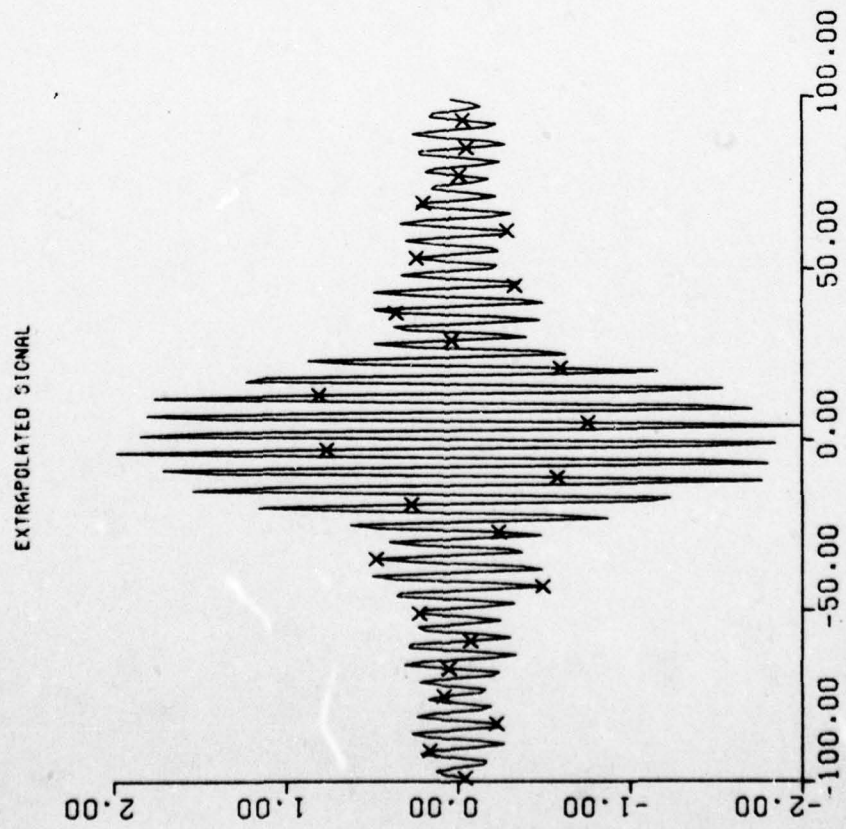


Fig. 5c: Signal Extrapolated by Papoulis' Iterative Algorithm

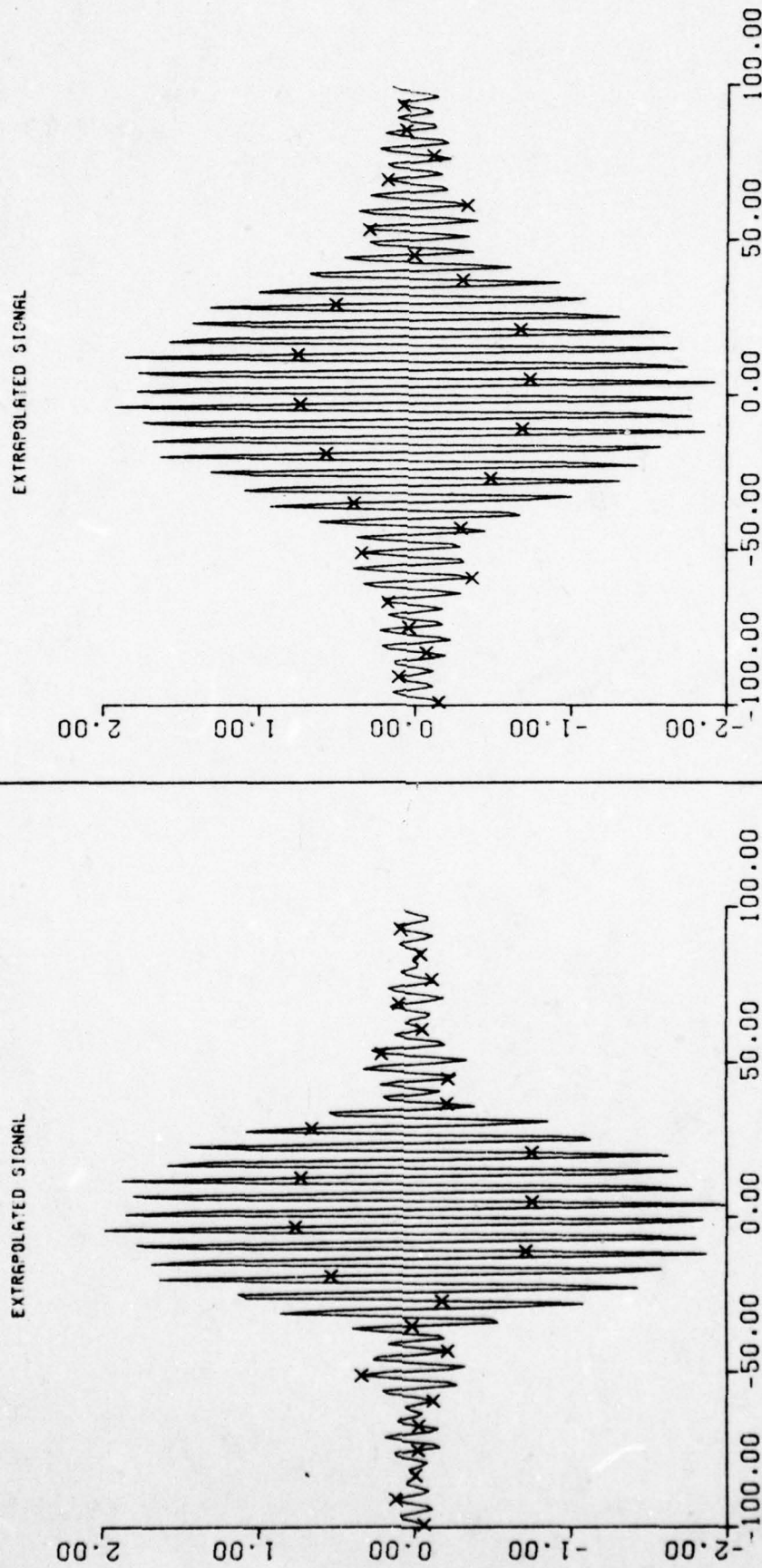


Fig. 5e: Signal Extrapolated After Adding a Stabilizing Diagonal Term to Matrix  $E_c$       Fig. 5f: Signal Extrapolated by Conjugate Gradient Algorithm

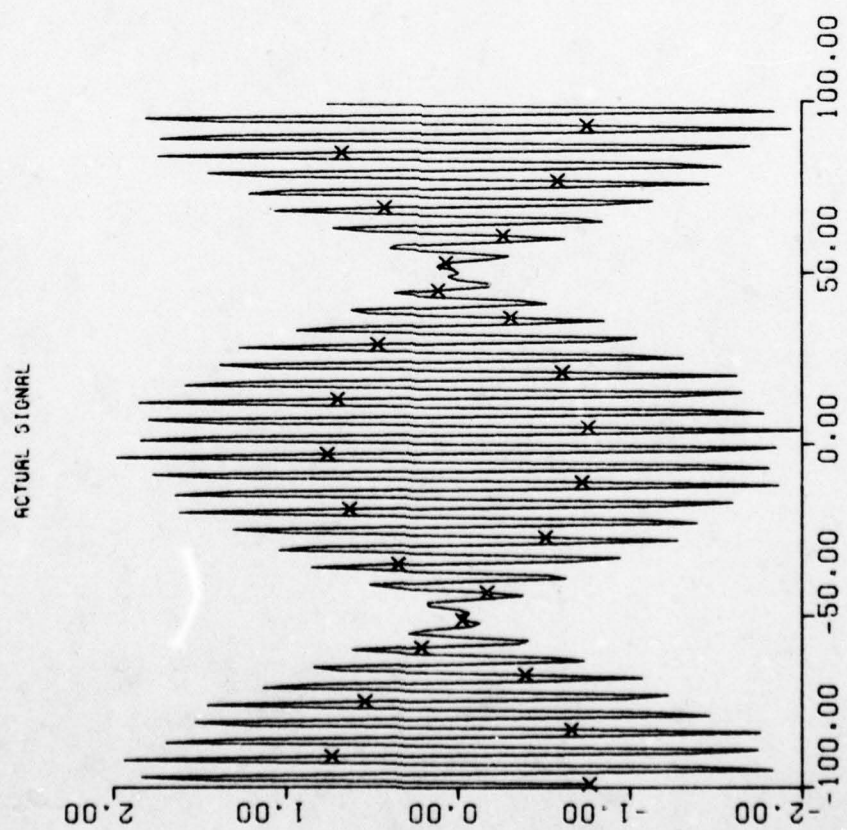
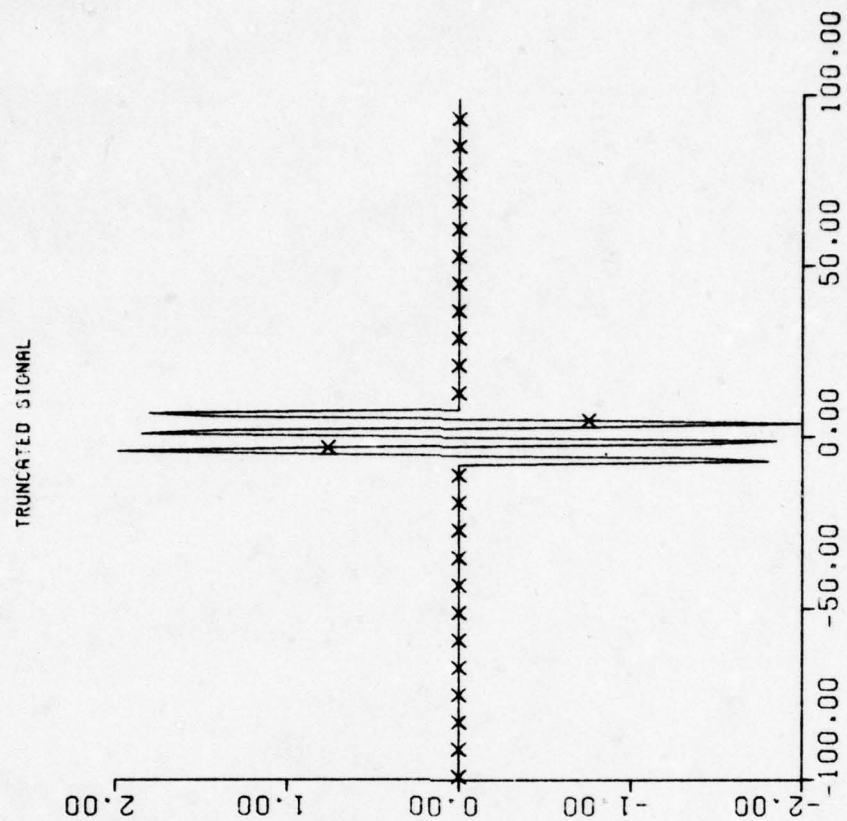


Fig. 6b: Given Observations (17 Samples)

Fig. 6a: Original Signal

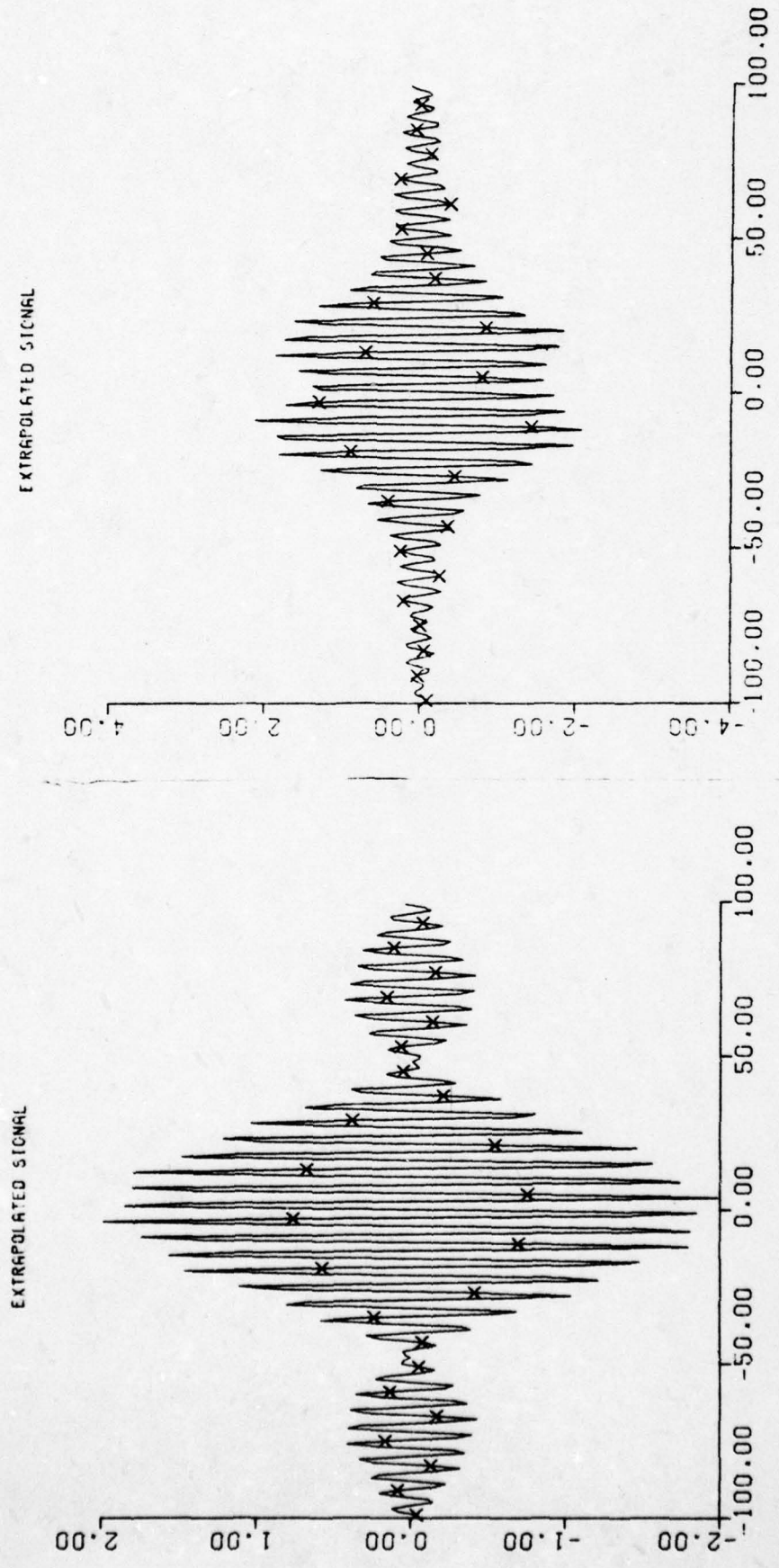


Fig. 3c: Signal Extrapolated by Papoulis' Iterative Algorithm      Fig. 3d: Signal Extrapolated Via Matrix  $E_c$

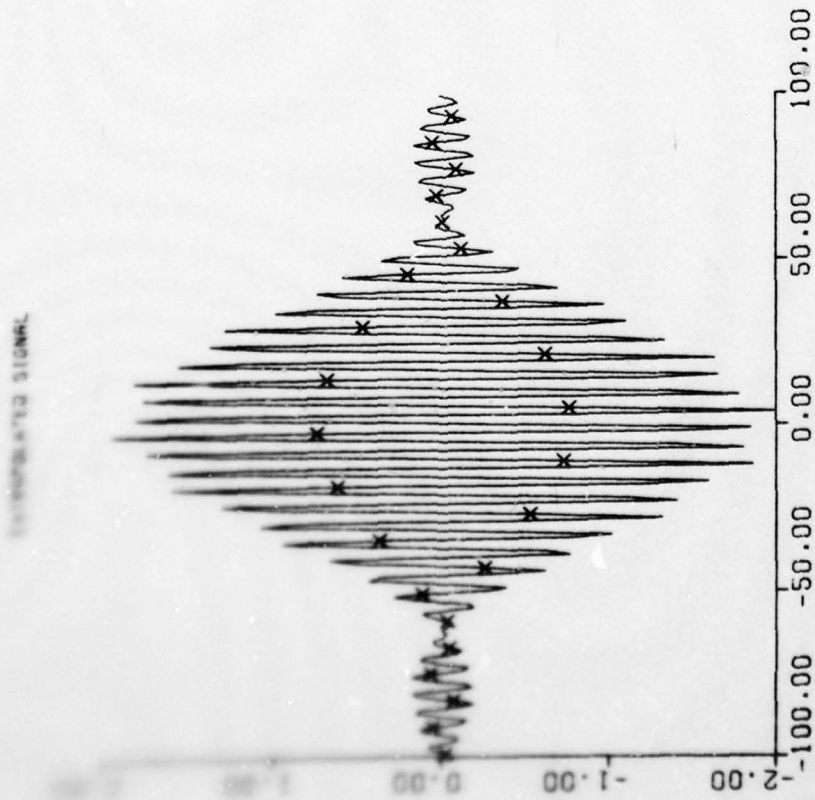
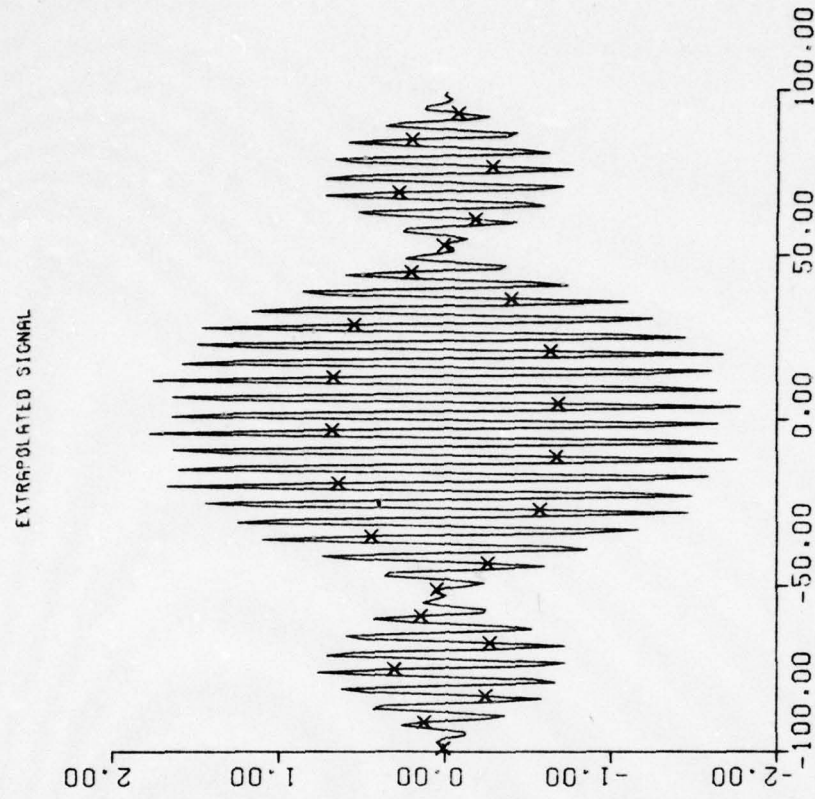


Fig. 6f: Signal Extrapolated by Conjugate Gradient Algorithm

Fig. 6e: Signal Extrapolated After Adding a Stabilizing Diagonal Term to Matrix  $E_c$

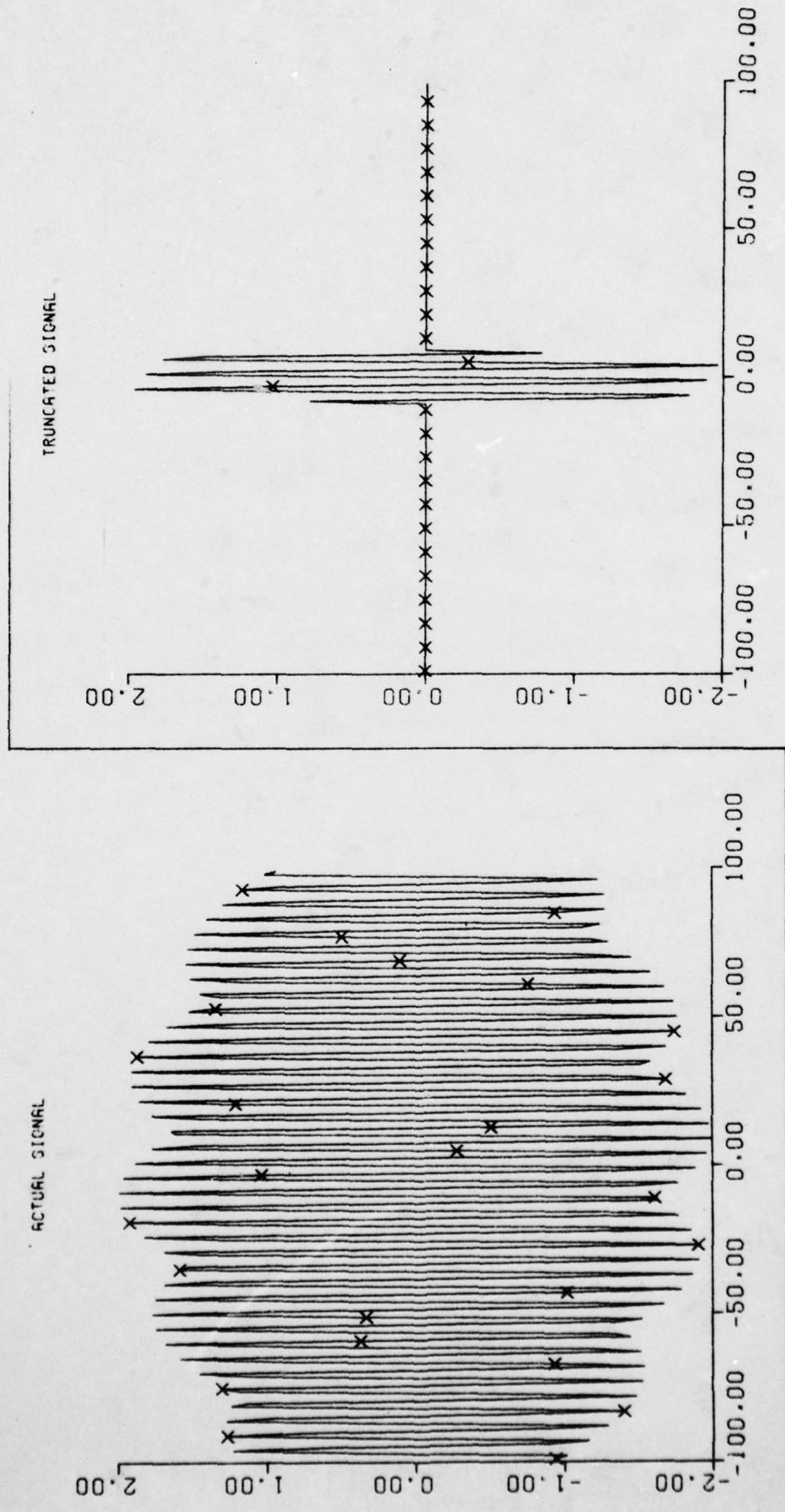


Fig. 7a: Original Signal

Fig. 7b: Given Observations (17 Samples)

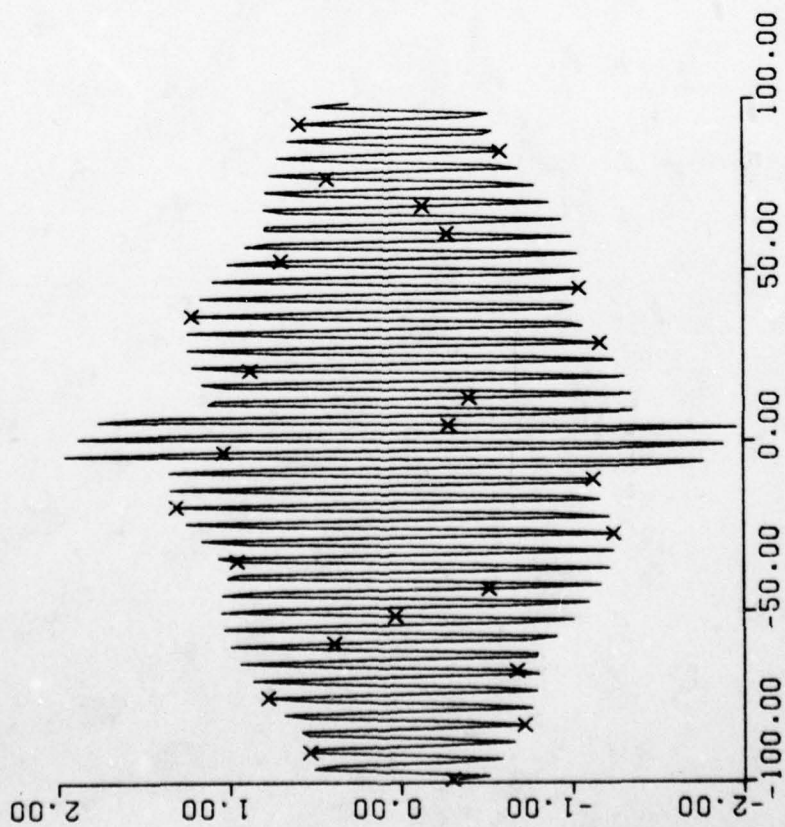
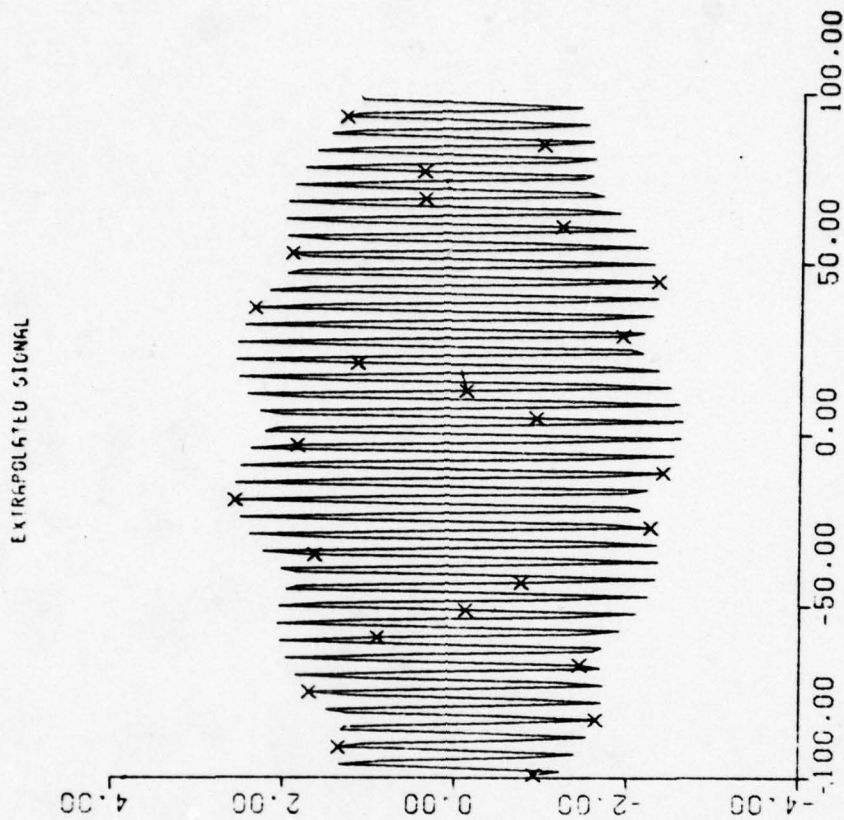


Fig. 7c: Signal Extrapolated by Papoulis' Iterative Algorithm      Fig. 7d: Signal Extrapolated Via Matrix  $E_c$

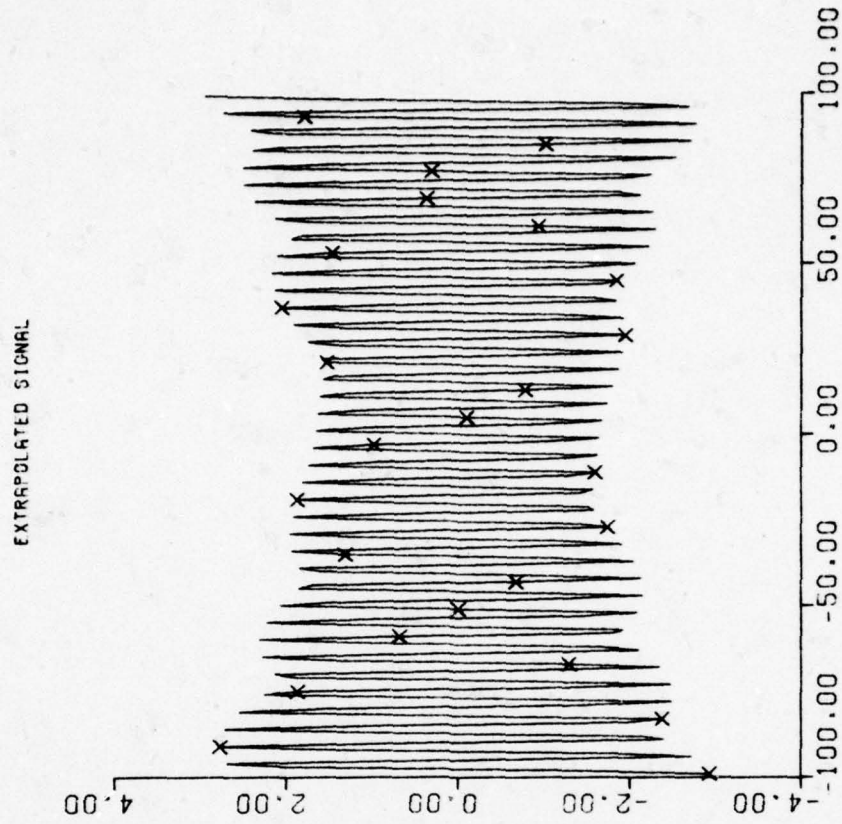


Fig. 7f: Signal Extrapolated by Conjugate Gradient Algorithm

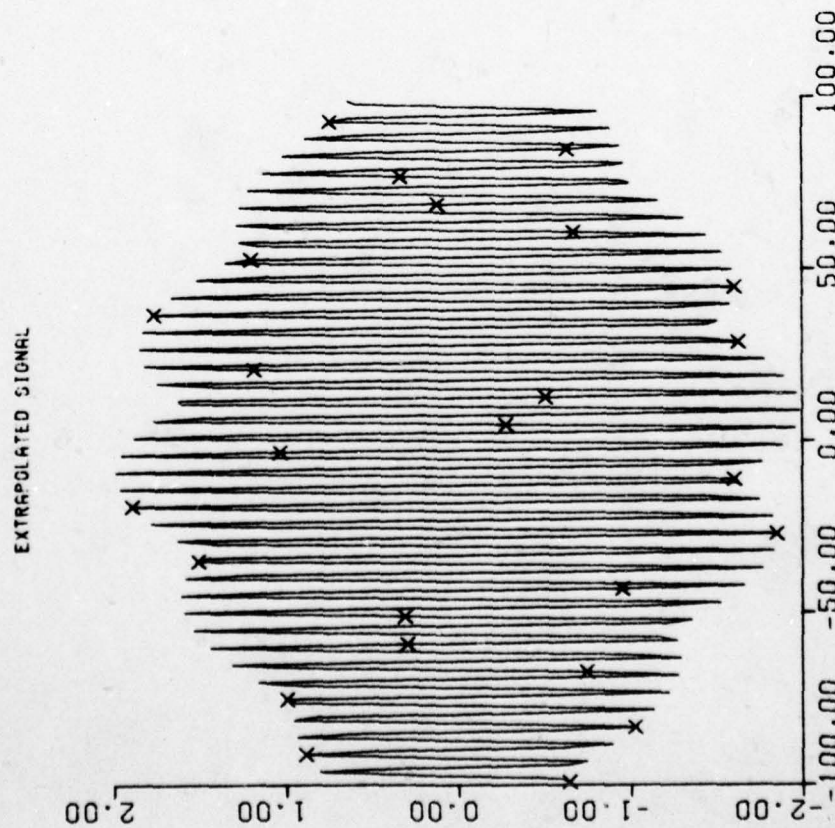


Fig. 7e: Signal Extrapolated After Adding a Stabilizing Diagonal Term to Matrix  $E_c$

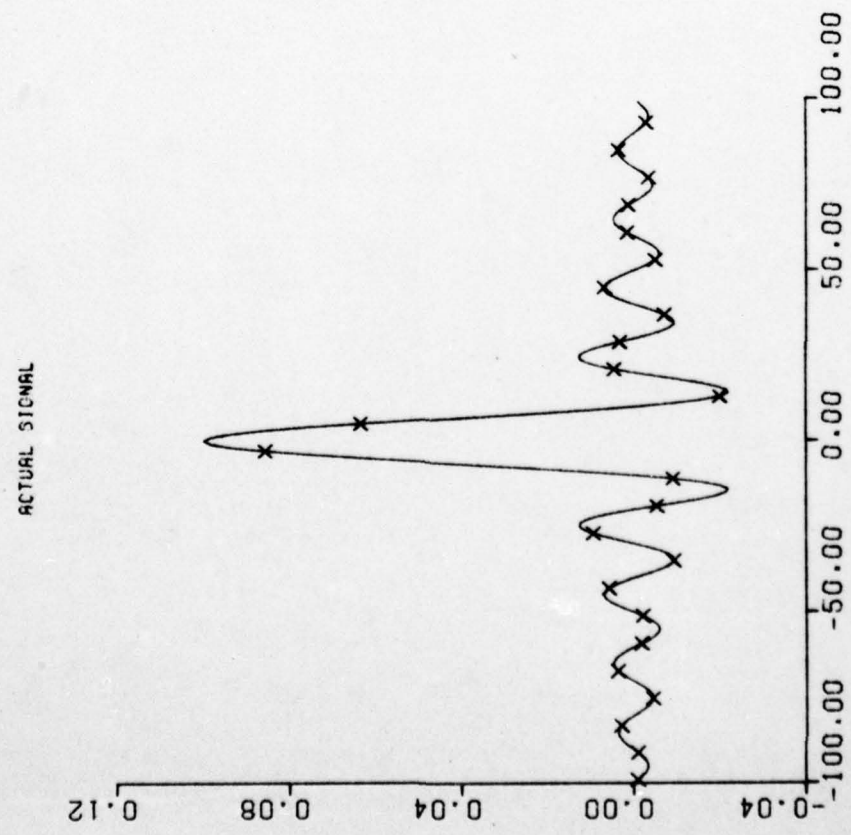


Fig. 8a: Original Signal

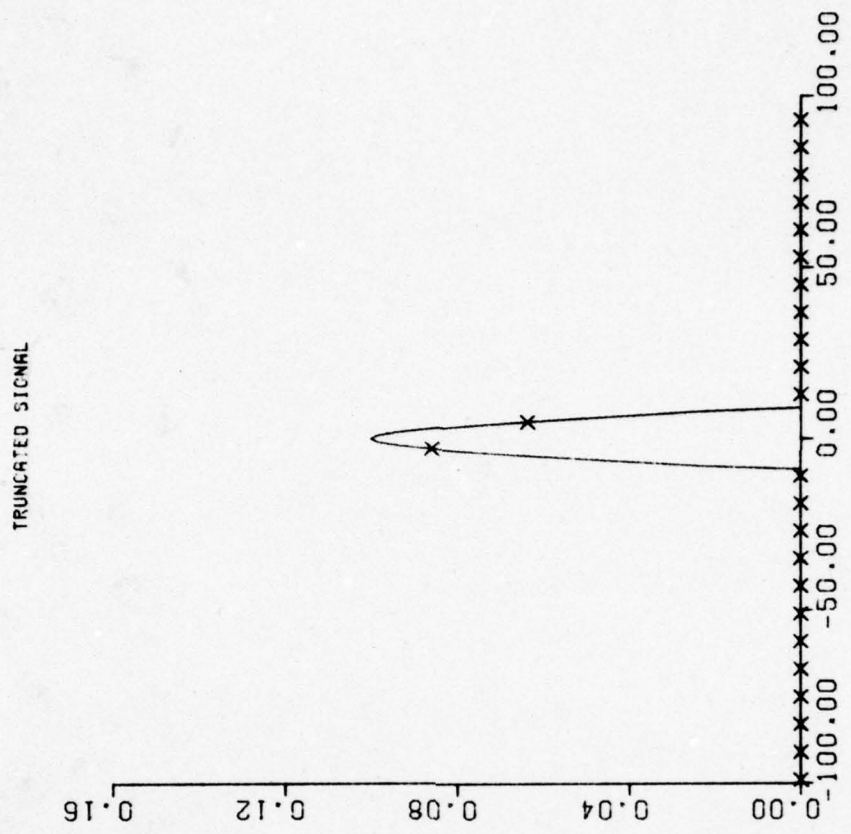


Fig. 8b: Given Observations (17 Samples)

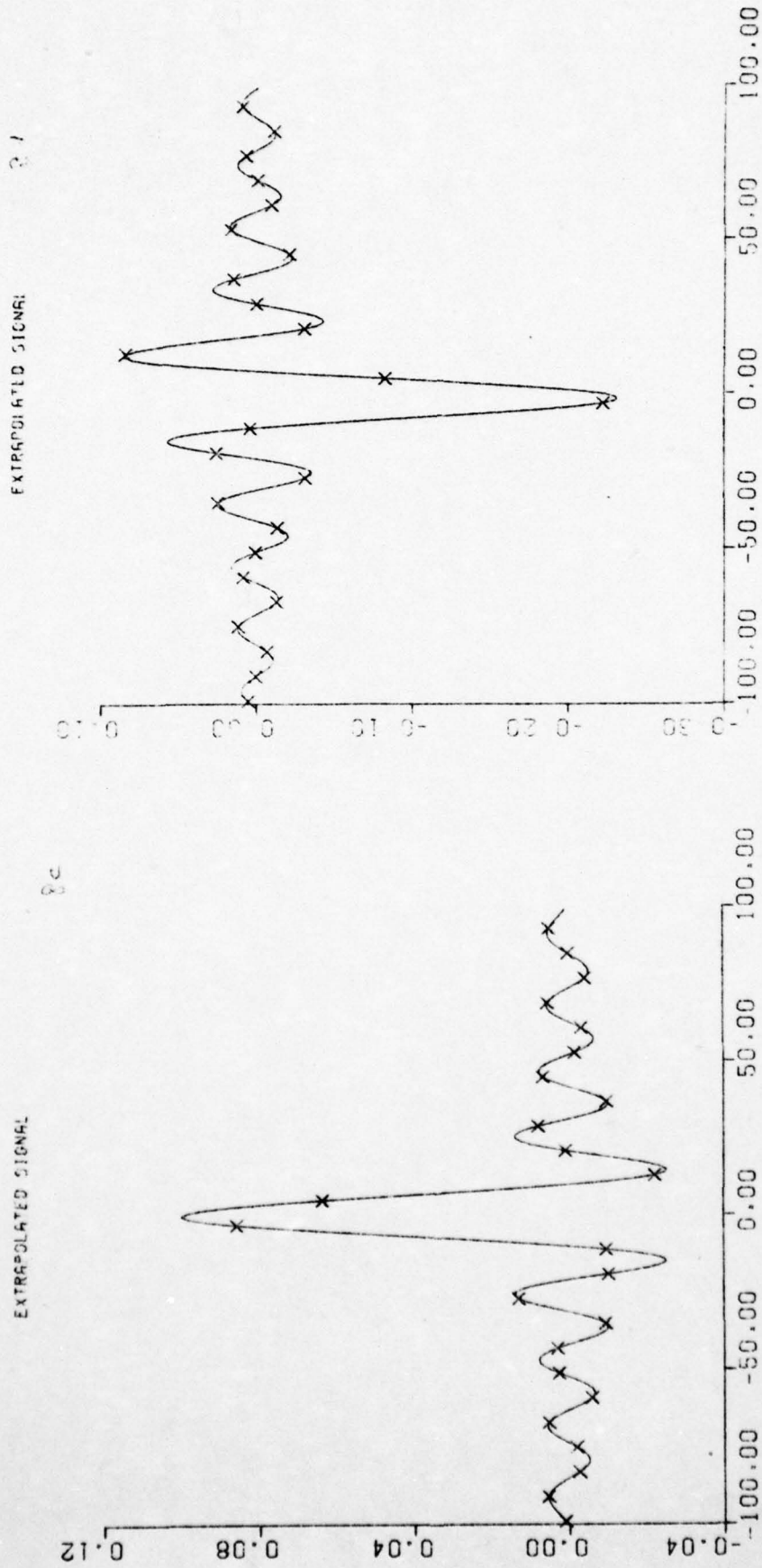


Fig. 8c: Signal Extrapolated by Papoulis' Iterative Algorithm      Fig. 8d: Signal Extrapolated Via Matrix  $E_c$

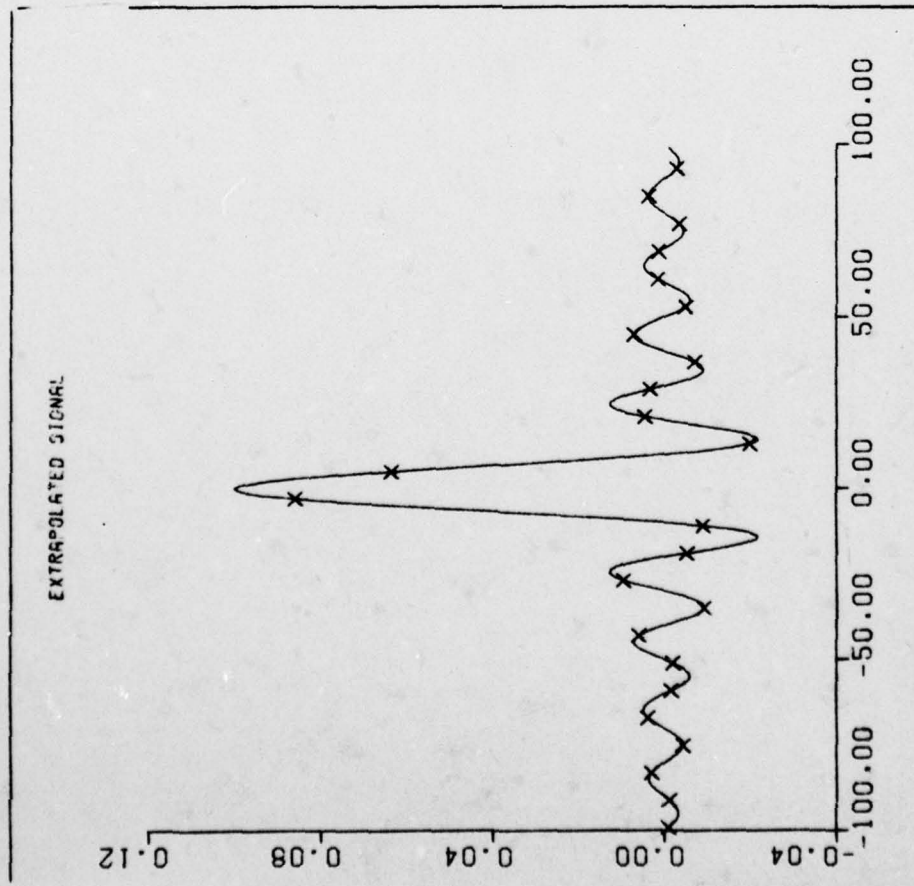


Fig. 8e: Signal Extrapolated After Adding a Stabilizing Diagonal Term to Matrix  $E_c$

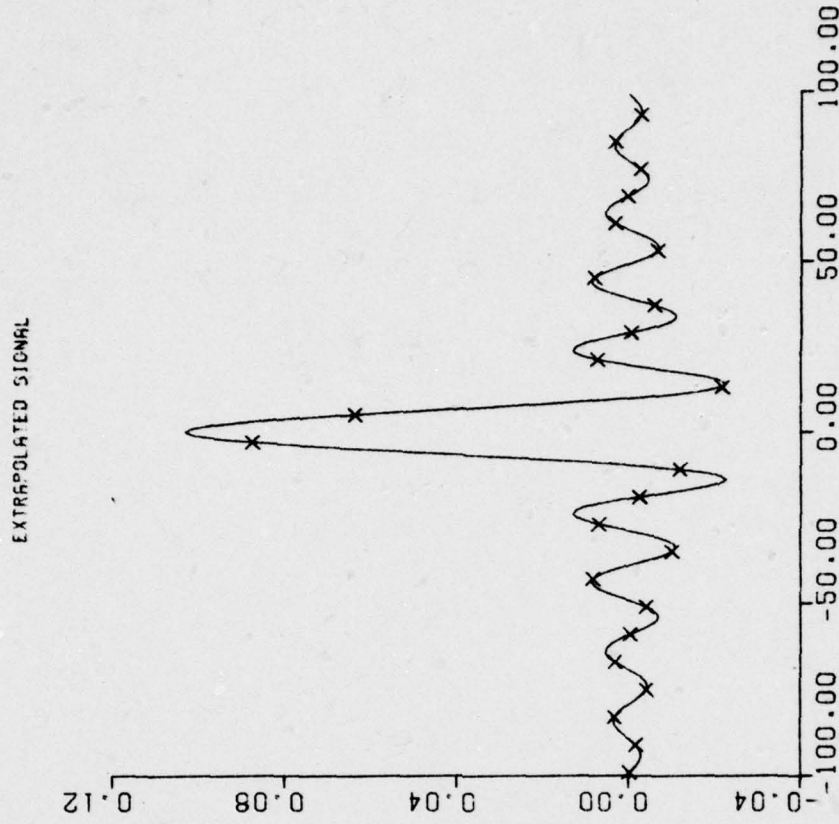


Fig. 8f: Signal Extrapolated by Conjugate Gradient Algorithm

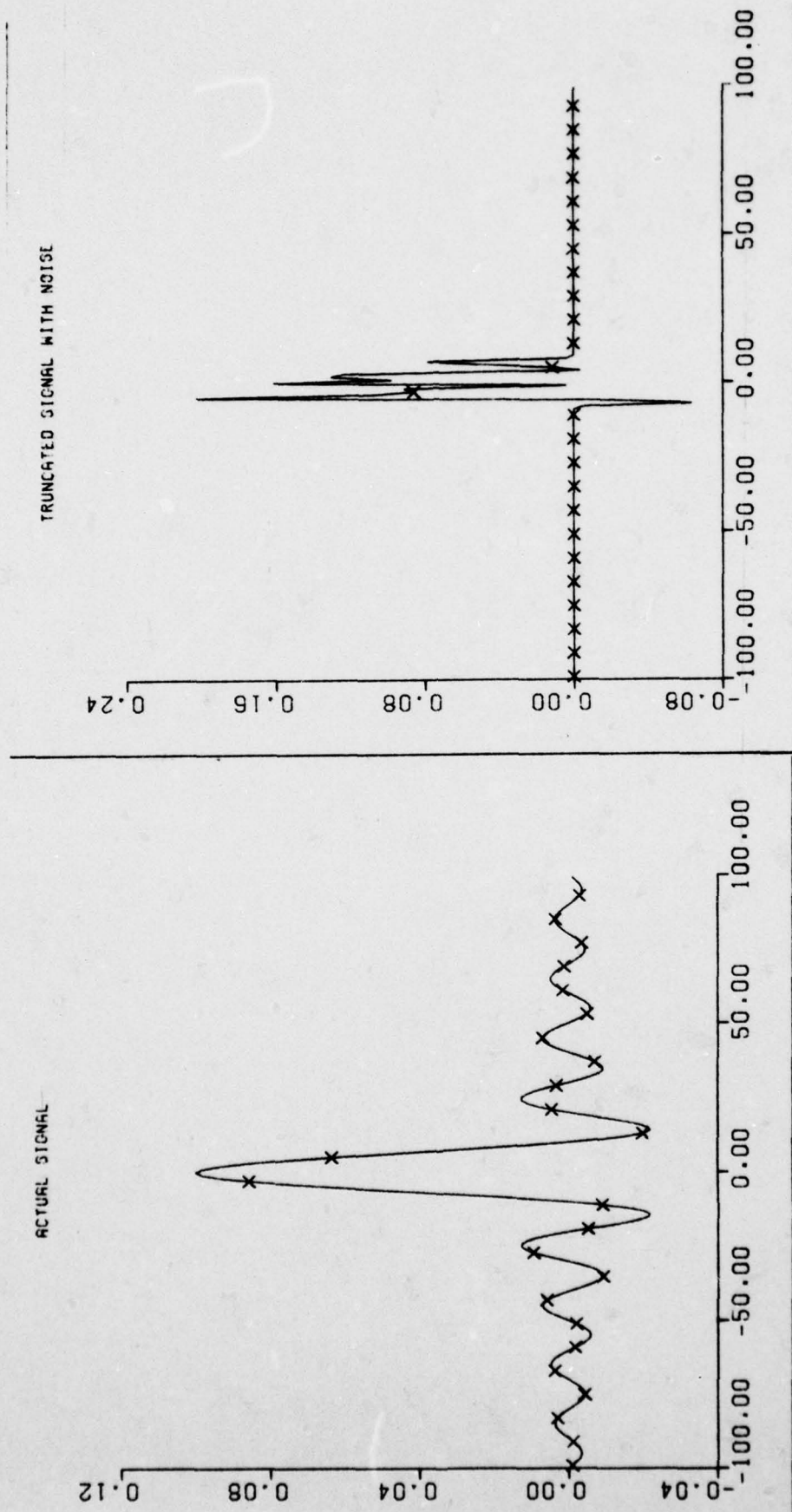


Fig. 9a: Original Signal

Fig. 9b: Given Observations (17 Samples)

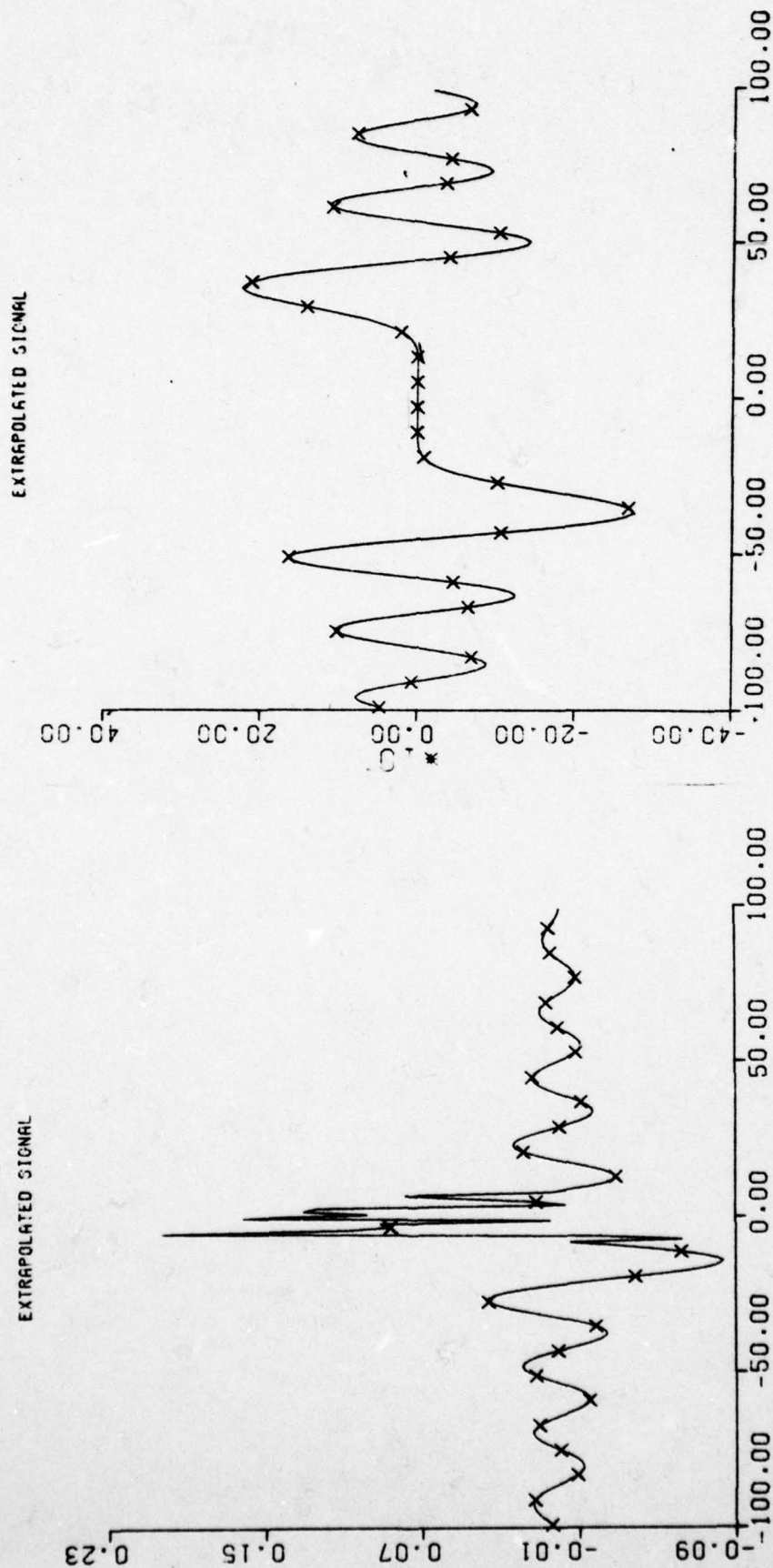


Fig. 9c: Signal Extrapolated by Papoulis' Iterative Algorithm Fig. 9d: Signal Extrapolated Via Matrix  $E_c$

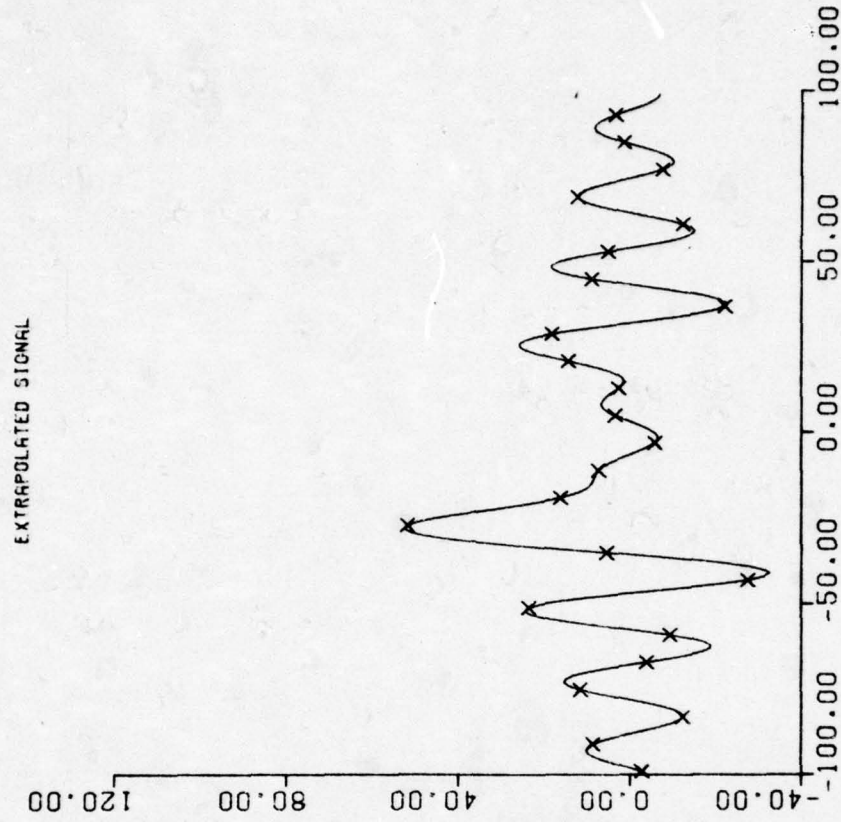


Fig. 9e: Signal Extrapolated After Adding a Stabilizing Diagonal Term to Matrix  $E_c$

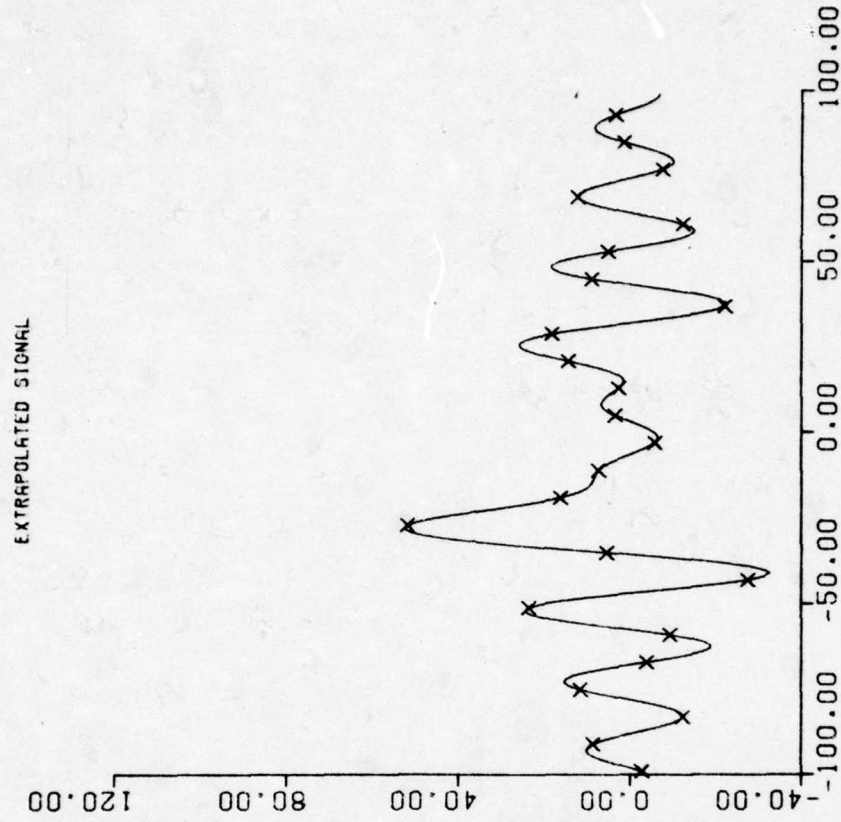


Fig. 9f: Signal Extrapolated by Conjugate Gradient Algorithm

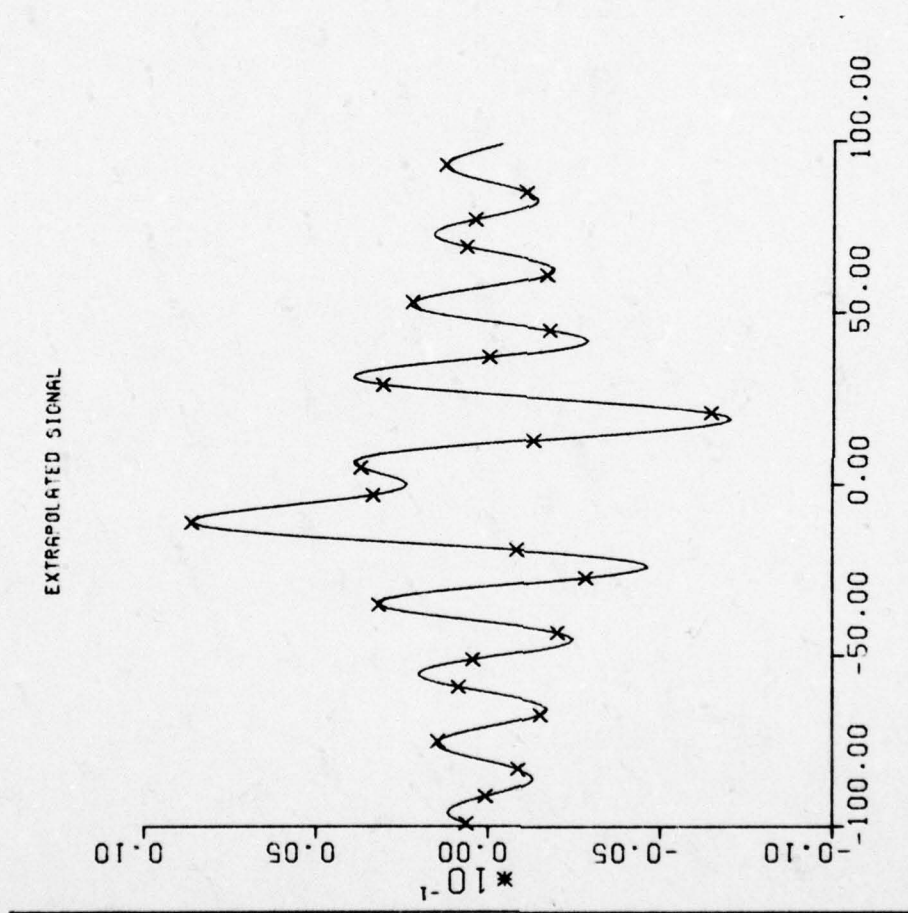
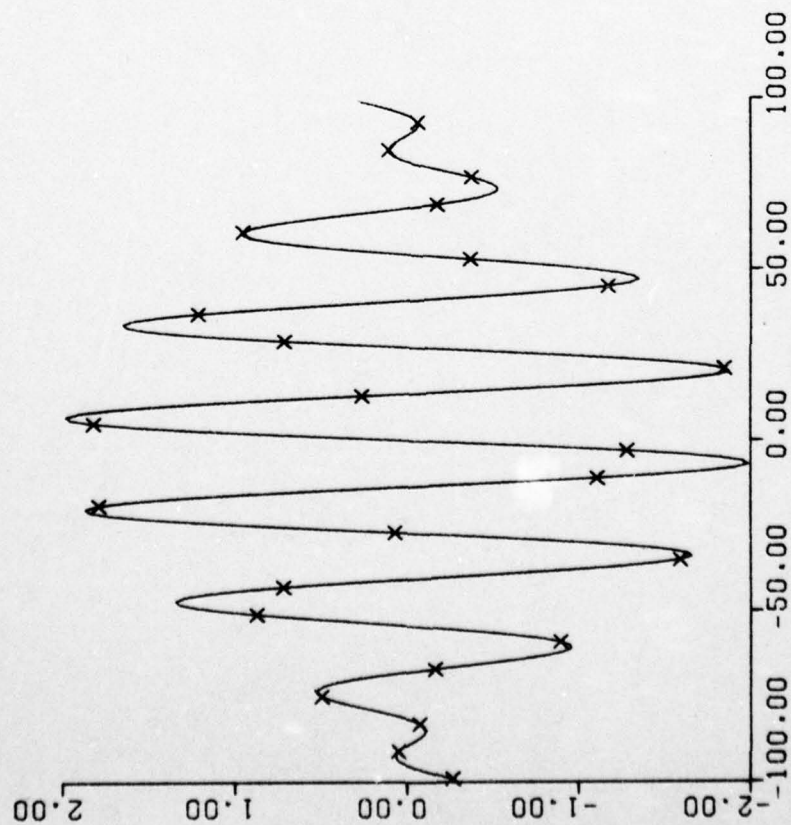


Fig. 9g: Signal Extrapolated by M.S. Extrapolation Filter

ACTUAL SIGNAL



TRUNCATED SIGNAL WITH NOISE

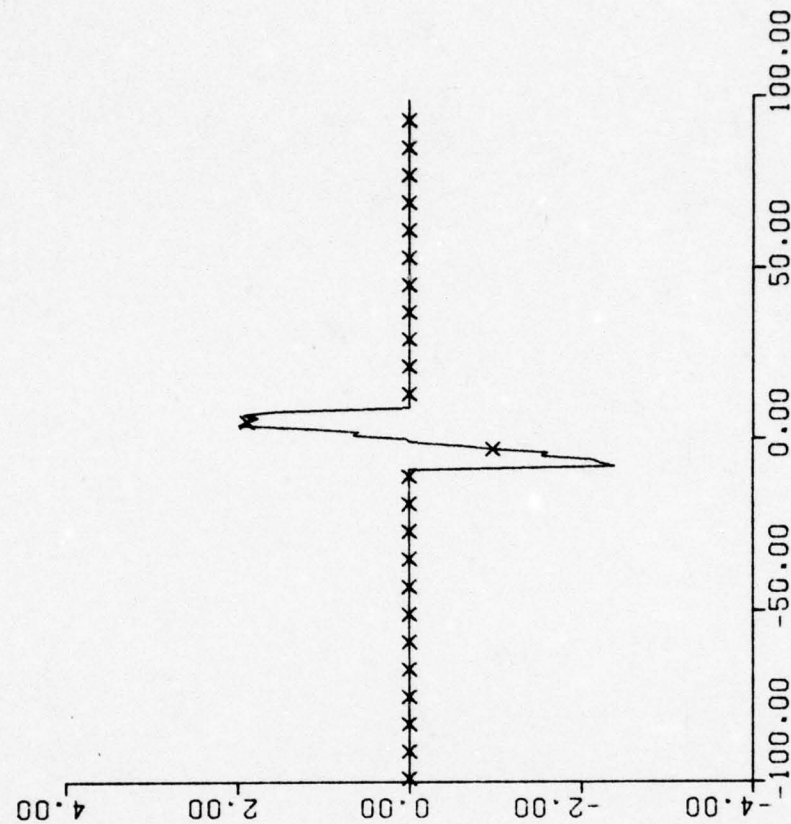


Fig. 10a: Original Signal

Fig. 10b: Given Observations (17 Samples)

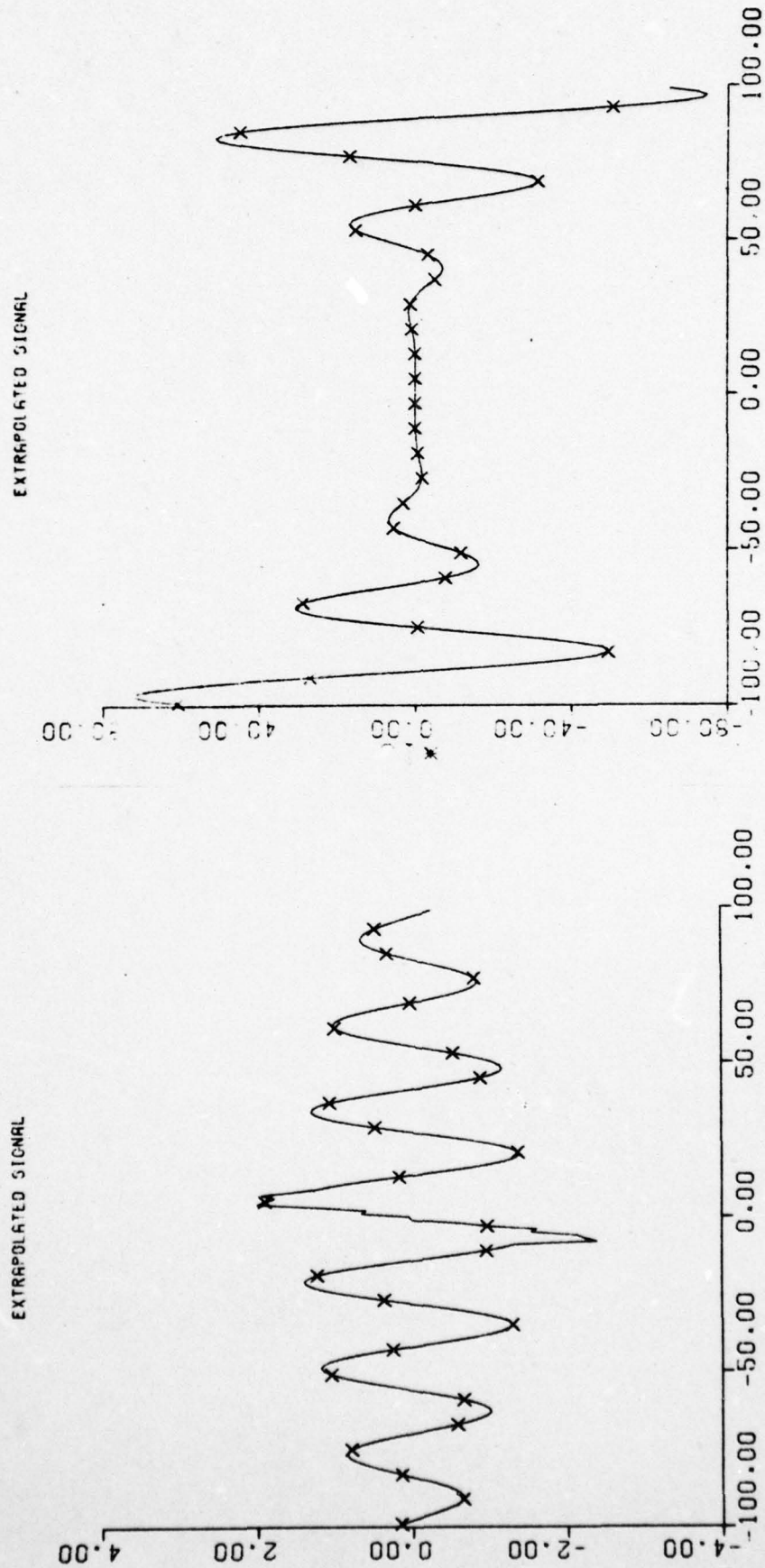


Fig. 10c: Signal Extrapolated by Papoulis' Iterative Algorithm      Fig. 10d: Signal Extrapolated Via Matrix  $E_c$

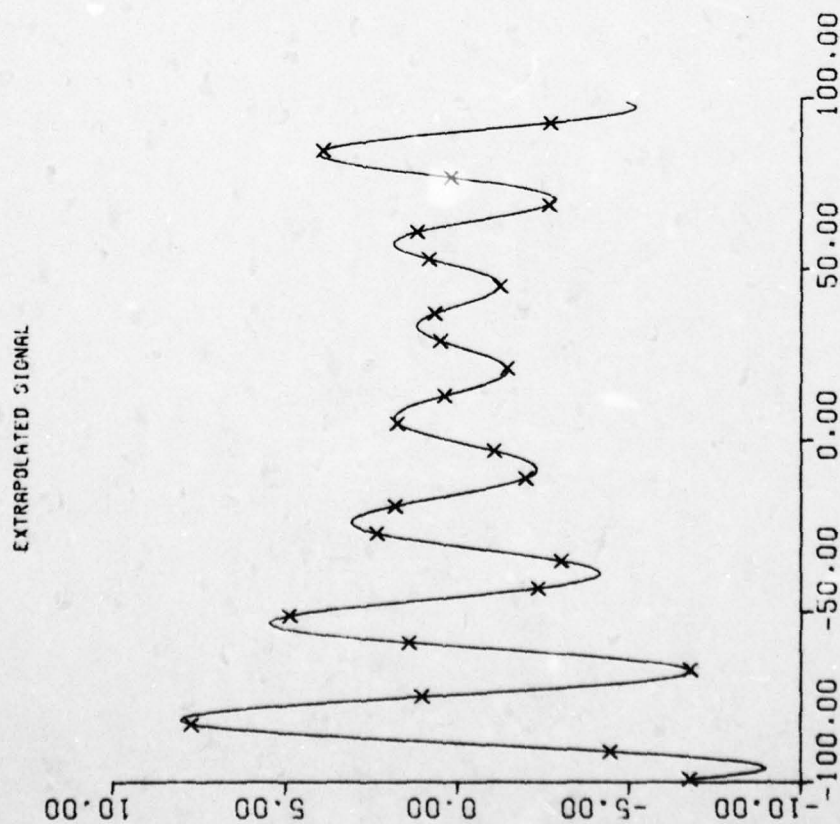
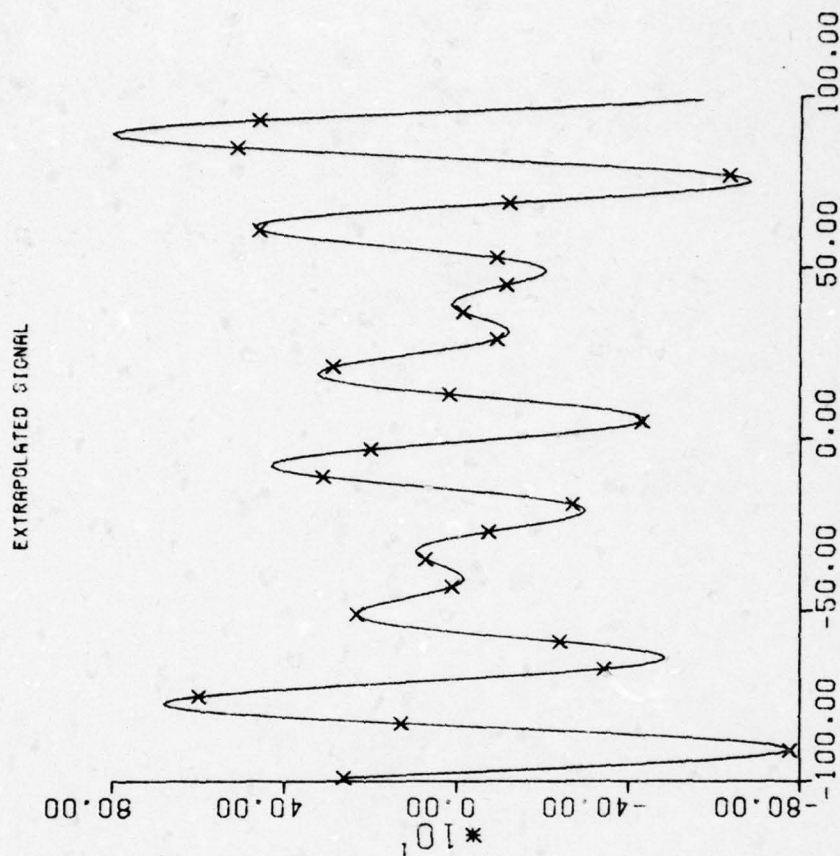


Fig. 10f: Signal Extrapolated by Conjugate Gradient Algorithm

Fig. 10e: Signal Extrapolated After Adding a Stabilizing Diagonal Term to Matrix  $E_c$

AD-A066 139

CALIFORNIA UNIV DAVIS SIGNAL AND IMAGE PROCESSING LAB F/G 9/3  
EXTRAPOLATION AND SPECTRAL ESTIMATION TECHNIQUES FOR DISCRETE T--ETC(U)  
NOV 78 A K JAIN, S RANGANATH DAAG29-78-A-0206

UNCLASSIFIED

SIPL-78-1

ARO-16222.1-EL

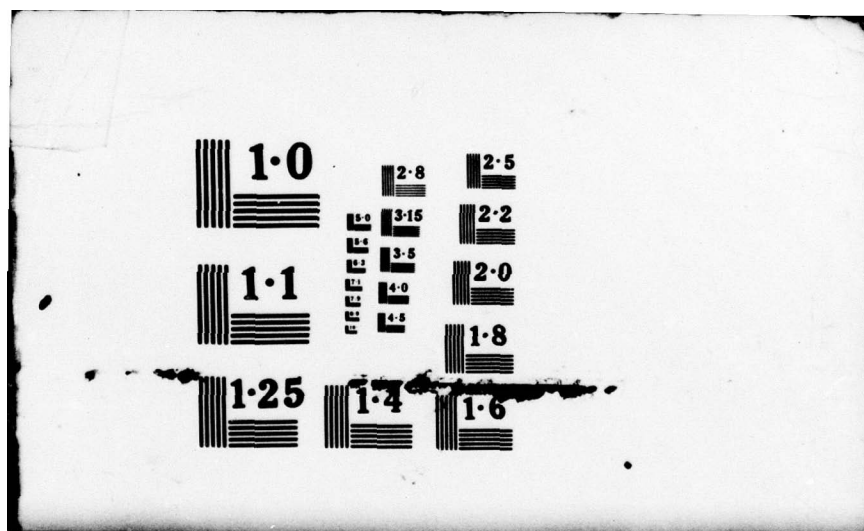
NL

2 OF 2  
AD A  
066139



END  
DATE  
FILMED

5-79  
DDC



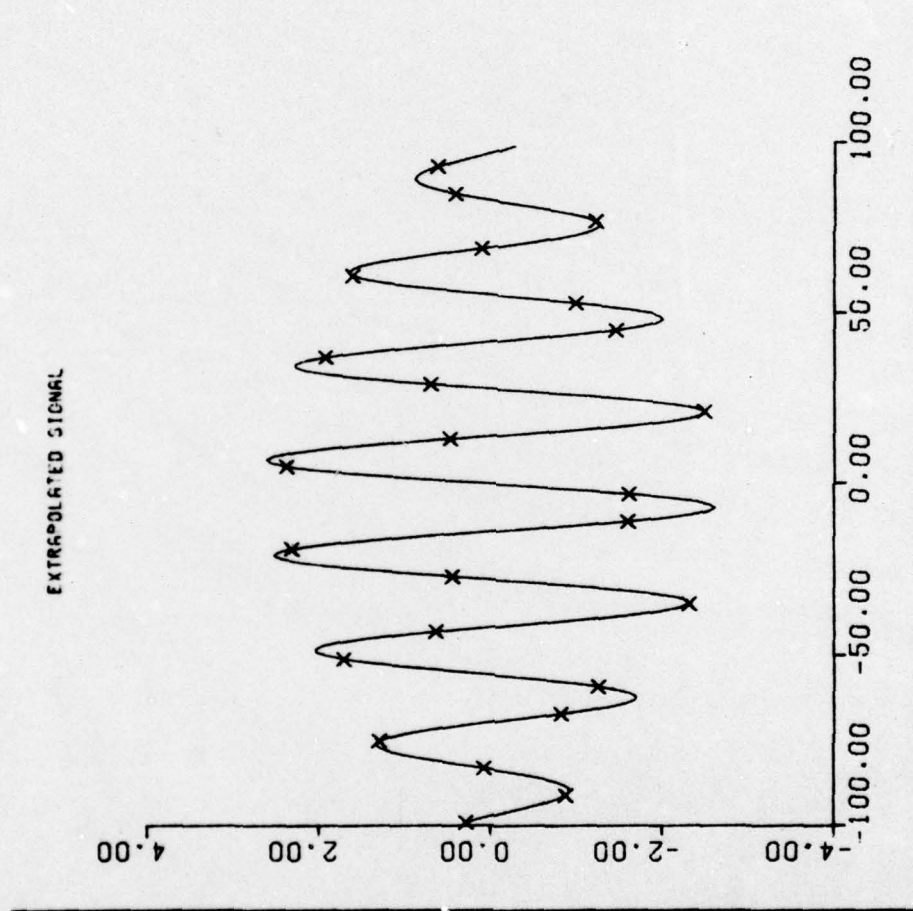


Fig. 10g: Signal Extrapolated by M.S. Extrapolation Filter

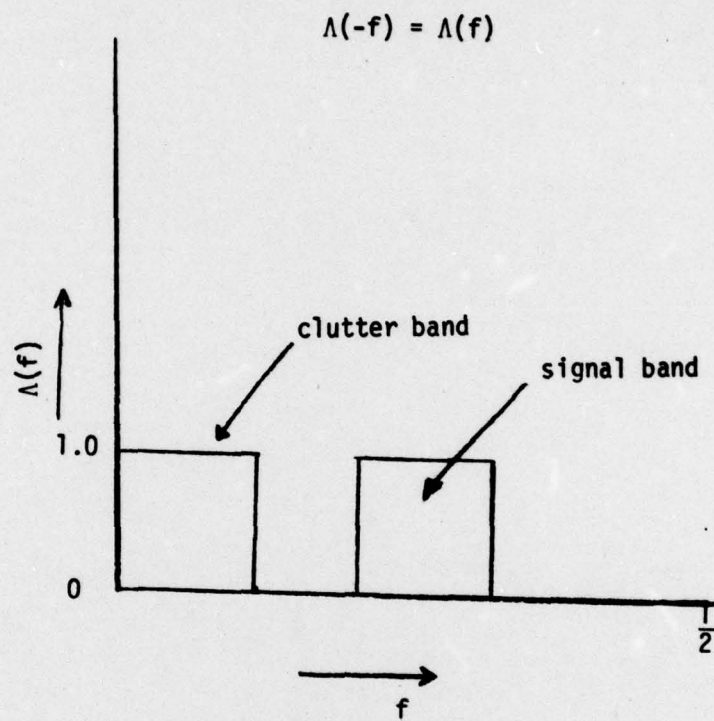


Fig. 11:  $\Lambda(f)$  vs.  $f$

## IX. CONCLUSIONS

The problem considered here was to effectively discriminate the signal from the interfering clutter, based on a small number of observation samples. Conventional techniques like the DFT and Maximum Entropy Methods are seen to yield poor results.

Given that the signal and clutter are band limited in mutually exclusive bands, a good method of improving resolution is to extrapolate the signal outside the observation interval and then estimate the spectrum. Though continuous time band limited signals can be extrapolated exactly outside the observation interval, this is not possible in the discrete time case. In fact, in this report we have proved that the discrete extension of the continuous algorithms leads to an extrapolated signal which is optimum in a minimum norm least squares sense.

We have introduced some new algorithms for signal extrapolation, viz.;

- a) Conjugate Gradient Algorithm
- b) Mean Square Extrapolation Filter
- c) Recursive Extrapolation Filter
- d) An Extrapolation Algorithm via Discrete PSWFs.

The Conjugate Gradient method is an iterative technique which has a rapid initial rate of convergence and hence has an obvious advantage over Papoulis' algorithm which has a linear rate of convergence. The Mean Square extrapolation filter is a non-iterative method that takes noise statistics into account and simultaneously filters the clutter from the signal. Cadzow's one shot method is seen to be a special case of this filter.

Further experiments have to be performed for the Recursive Extrapolation filter (which also considers noise statistics) and for extrapolation via Discrete PSWFs. In practice, extrapolation is achieved only to a limited extent beyond the observation interval and depends on the signal bandwidth uncertainty. The larger this bandwidth uncertainty, the smaller the length to which the signal can be extrapolated. Although, in the absence of noise, all algorithms yield a minimum norm least squares solution, their implementations are different and lead to different truncation errors. Such error analysis and other numerical features for the extrapolation problem need further work. Our experiments show that whatever the uncertainty of the signal bandwidth might be, extrapolation of the signal followed by a spectral estimation improves the spectral estimate.

Table 1

	METHOD	COMPUTATIONAL COMPLEXITY	GENERAL COMMENTS
1	MAX. ENTROPY or Autoregressive (Burg, Parzen & others)	pxp Toeplitz Eqns. + FFT operation $\approx 3p^2 + O(N \log N)$ ; $p \ll N$	<ol style="list-style-type: none"> <li>1. Simple linear Eqns., easy to implement.</li> <li>2. Good results for all-pole spectra.</li> <li>3. Fails in the presence of noise and clutter-unless a large number of observation samples are available.</li> <li>4. Order of the model difficult to select.</li> <li>5. Performance improved by applying it on extrapolated signal.</li> </ol>
2	Continuous PSWF (Slepian et al.)	Very large. Functions extremely difficult to calculate	<ol style="list-style-type: none"> <li>1. For extrapolation of bandlimited, continuous signals; existence guaranteed.</li> <li>2. Extremely difficult to implement.</li> <li>3. Noise sensitive</li> <li>4. Useful in establishing existence, uniqueness &amp; other properties.</li> </ol>
3	Iterative Extrapolation (Papoulis)	$O(4N \log_2 N)$ real operations per iteration	<ol style="list-style-type: none"> <li>1. Easy to implement</li> <li>2. Is a gradient algorithm with linear convergence. Requires a large number of iterations and FFT operations at each iteration.</li> <li>3. Does not take into account noise statistics.</li> </ol>
4	Extrapolation Matrix, $E_\infty$ (Sabri and Steenaart)	If observed data = $2M+1$ and extrapolated length = $N$ then $\sim O(N^3)$ to invert $N \times N$ Matrix $+1/2(2M+1)(N-2M-1)$ operations to find extrapolated signal	<ol style="list-style-type: none"> <li>1. <math>E_\infty</math> does not exist. <math>E_N</math> exists.</li> <li>2. A large (<math>N \times N</math>) ill-conditioned matrix has to be inverted.</li> <li>3. Noise sensitive. Can be stabilized by adding a diagonal term to <math>G</math>.</li> <li>4. Noise statistics not considered.</li> </ol>
5	$E_c$ (Cadzow)	If observed data = $m$ pts extrapolated to $N$ pts, $O(3m^2 + m^2 + mN)$	<ol style="list-style-type: none"> <li>1. Easy to implement, if <math>E_c</math> is stable.</li> <li>2. An ill-conditioned matrix (<math>m \times m</math> Toeplitz) has to be inverted.</li> <li>3. Noise sensitive. Can be stabilized.</li> <li>4. Noise statistics not considered.</li> </ol>
6	Conjugate Gradient (Jain & Ranganath)	$\sim O(2mN)$ operations per iteration	<ol style="list-style-type: none"> <li>1. Easy to implement</li> <li>2. Extremely rapid initial rate of convergence</li> <li>3. Small number of iteration required in practice</li> <li>4. Noise sensitive, but can be stabilized</li> <li>5. Noise statistics not considered</li> </ol>

Table 1, continued

	METHOD	COMPUTATIONAL COMPLEXITY	GENERAL COMMENTS
7	Mean Square Extrapolation Filter (Jain & Ranganath)	$\sim (3m^2 + m^2 + mN)$ operations. $mN$ operation once Filter gain has been computed.	<ol style="list-style-type: none"> <li>1. Easy to implement</li> <li>2. An <math>m \times m</math> Toeplitz matrix has to be inverted.</li> <li>3. Takes Noise statistics into account</li> <li>4. Reduces to <math>E_c</math> in the noise free case.</li> </ol>
8	Recursive Extrapolation (Jain & Ranganath)	$N$ operations per data point, if gains are pre-computed.	<ol style="list-style-type: none"> <li>1. Easily implemented, updates extrapolated estimate as new data arrives.</li> <li>2. Takes noise statistics into account.</li> </ol>
9	Discrete PSWF Singular Value Expansion (Papoulis, Jain & Ranganath)	Requires solving for eigenvalues and eigenvectors of a $(2M+1) \times (2M+1)$ Toeplitz matrix, a low pass filtering operation and a finite series expansion.	<ol style="list-style-type: none"> <li>1. More difficult to implement than (7) or (8). Easy to implement once the eigenvectors have been computed</li> <li>2. Noise statistics not considered.</li> <li>3. Accuracy depends on the accuracy of eigenvalues and eigenvectors.</li> </ol>

## X. BIBLIOGRAPHY

1. R. B. Blackman and J. W. Tukey, The Measurement of Power Spectra, Dover Publications, Inc., New York, 1958.
2. J. Makhoul, "Linear Prediction: A Tutorial Review," Proc. IEEE, Vol. 63, pp. 561-80, April 1975.
3. J. P. Burg, "Maximum Entropy Spectral Analysis," Ph.D. Thesis, Stanford University, 1975.
4. E. Parzen, "Multiple Time Series Modeling," in Multivariate Analysis-II, (P. R. Krishnaiah, ed.), Academic Press, New York, pp. 389-409, 1969.
5. H. R. Radoski, D. F. Fougere and E. J. Zowalick, "A Comparison of Power Spectral Estimates and Applications of the Maximum Entropy Method," J. of Geophys. Res., Vol. 80, pp. 619-625, 1975.
6. M. Morf, et. al., "Efficient Solution of Covariance Equations for Linear Prediction," IEEE Trans. ASSP., Vol. ASSP-26, pp. 429-433, Oct. 1977.
7. T. J. Ulrych and R. W. Clayton, "Time Series Modelling and Maximum Entropy," Phys. Earth Planet. Inter., Vol. 12, pp. 188-200, 1976.
8. A. Papoulis, "A New Algorithm in Spectral Analysis and Band-limited Extrapolation," IEEE Trans. on Circuits and Systems, Vol. CAS-22, #9, pp. 735-742, Sept. 1975.
9. R. W. Gerchberg, "Super Resolution Through Error Energy Reduction," Optica Acta, Vol. 21, No. 9, pp. 709-720, 1974.
10. M. S. Sabri and W. Steenaart, "An Approach to Bandlimited Signal Extrapolation: The Extrapolation Matrix," IEEE Trans. on Circuits and Systems, Vol. CAS-25, #2, pp. 74-78, Feb. 1978.

11. J. A. Cadzow, "Improved Spectral Estimation from Incomplete Sampled Data Observations," Presented at the RADC Spectrum Estimation Workshop, Rome, New York, May 24-26, 1978
12. D. G. Luenberger, Introduction to Linear and Nonlinear Programming, Addison-Wesley, Reading, Mass., 1973.
13. E. S. Angel and A. K. Jain, "Restoration of Images Degraded by Spatially Varying Pointspread Functions by a Conjugate Gradient Method," Applied Optics, Vol. 17, #16, pp. 2186-2190, July 15, 1978.
14. D. S. Slepian, et. al., "Prolate Spheroidal Wave Functions, Fourier Analysis and Uncertainty Principle," Bell Syst. Tech. J., Vol. 40, #1, pp. 43-84, 1961.
15. E. Parzen, Time Series Analysis Papers Holden Day, San Francisco, pp. 940, 1967.
16. N. E. Nahi, Estimation Theory and Applications, John Wiley, New York, pp. 201-206, 1969.
17. A. K. Jain, "An Operator Factorization Method for Restoration of Blurred Images," IEEE Transactions on Computers, vol. C-26, #11, pp. 1061-1071, November 1977.
18. C. R. Rao and S. K. Mitra, Generalized Inverse of Matrices and Applications, Wiley, New York, 1971.
19. B. R. Frieden, "Band Unlimited Reconstruction of Optical Objects and Spectra," J. Opt. Soc. Am., vol. 57, No. 8, pp. 1013-1019, August 1967.
20. A. Papoulis and M. S. Bertram, "Digital Filtering and Prolate Functions," IEEE Trans. Circuit Theory, vol. CT-19, No. 6, pp. 674-681, November 1972.
21. G. H. Golub, "Least squares, singular values, and matrix approximations," Aplikace Matematiky, vol. 13, pp. 44-51, 1968.

22. A. K. Jain, Multidimensional Techniques in Computer Image Processing, book, (to appear).
23. M. M. Sondhi, "Image Restoration: The Removal of Spatially Invariant Degradations," Proc. IEEE, vol. 60, No. 7, pp. 842-853, July 1972.
24. Isaacson and H. B. Keller, Numerical Solution of Linear Systems, John Wiley & Sons, New York, pp. 75-76, 1966.